Technical Report HL-95-7
September 1995

**US Army Corps of Engineers**
Waterways Experiment Station

# Guidance on Generating Hydrologic Model Input from a Triangulated Irregular Network (TIN) Using a Curvature Based Filtering Technique to Extract TIN Vertices from Gridded Digital Elevation Models

*by Jeffrey D. Jorgeson*





Approved For Public Release; Distribution Is Unlimited



Prepared for Headquarters, U.S. Army Corps of Engineers

Technical Report HL-95-7
September 1995

# Guidance on Generating Hydrologic Model Input from a Triangulated Irregular Network (TIN) Using a Curvature Based Filtering Technique to Extract TIN Vertices from Gridded Digital Elevation Models

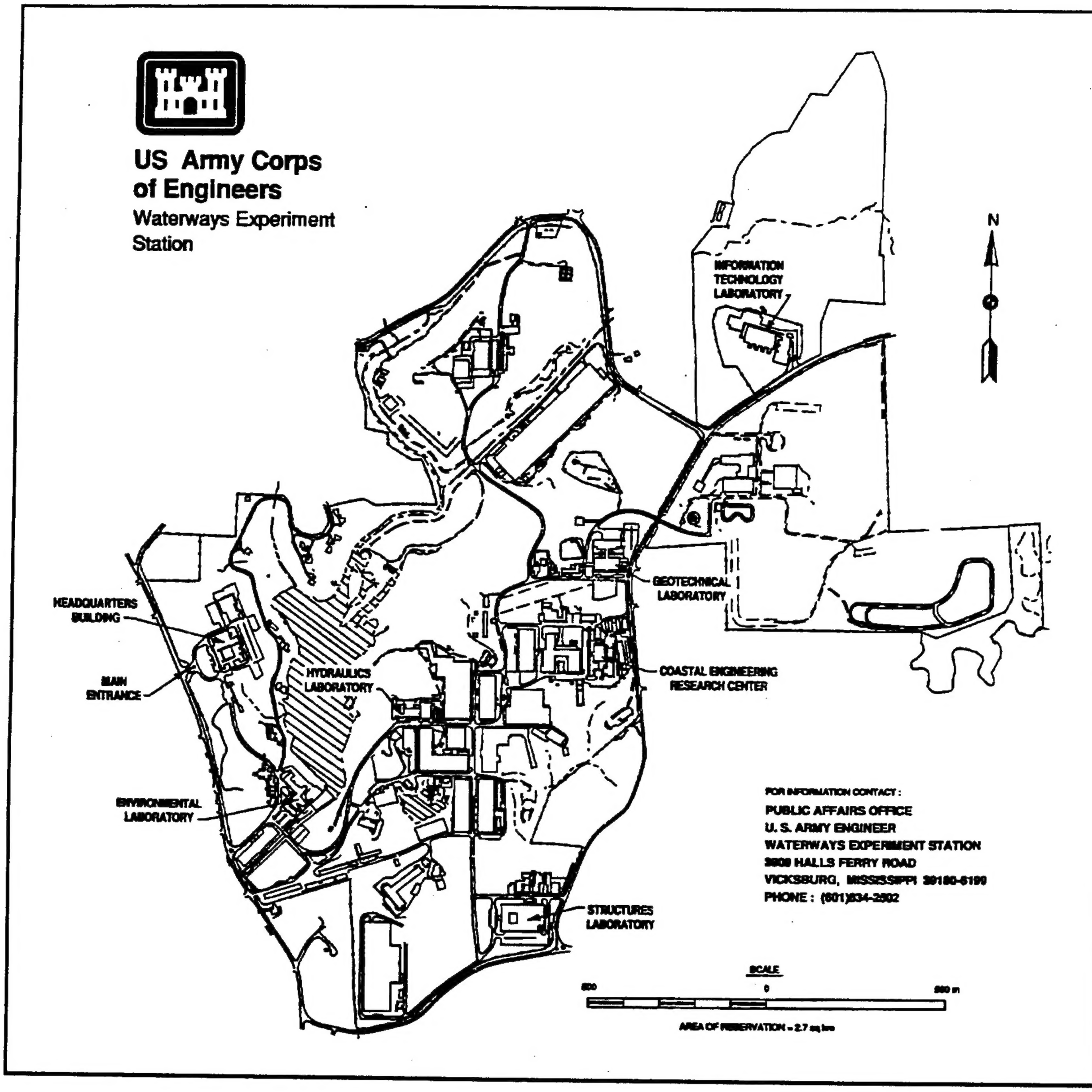
by Jeffrey D. Jorgeson

U.S. Army Corps of Engineers
Waterways Experiment Station
3909 Halls Ferry Road
Vicksburg, MS 39180-6199

Final report

Approved for public release; distribution is unlimited

Prepared for U.S. Army Corps of Engineers
Washington, DC 20314-1000




**Waterways Experiment Station Cataloging-in-Publication Data**

Jorgeson, Jeffrey D.

Guidance on generating hydrologic model input from a triangulated irregular network (TIN) using a curvature based filtering technique to extract TIN vertices from gridded digital elevation models / by Jeffrey D. Jorgeson ; prepared for U.S. Army Corps of Engineers.

214 p. : ill. ; 28 cm. — (Technical report ; HL-95-7)

1. Digital mapping. 2. Hydrologic models. 3. Triangulation. I. U.S. Army. Corps of Engineers. II. U.S. Army Engineer Waterways Experiment Station. III. Hydraulics Laboratory (U.S. Army Engineer Waterways Experiment Station) IV. Title. V. Series: Technical report (U.S. Army Engineer Waterways Experiment Station) ; HL-95-7.

TA7 W34 no.HL-95-7

TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

FIGURE Page

FIGURE Page

FIGURE Page

## PREFACE

This research was conducted at the Hydraulics Laboratory (HL) of the U.S. Army Engineer Waterways Experiment Station (WES) during the period of September 1994 through May 1995. The research was conducted and report prepared by Mr. Jeffrey D. Jorgeson, Watershed Systems Group, Hydro-Science Division, HL, as a Master's Thesis in the Department of Civil Engineering at Mississippi State University.

This report was prepared under the direct supervision of Mr. William D. Martin, Acting Chief, Hydro-Science Division; and under the general supervision of Mr. Robert F. Athow, Acting Assistant Director, HL; and Mr. Richard A. Sager, Acting Director, HL.

At the time of publication of this report, Director of WES was Dr. Robert W. Whalin, and Commander was COL Bruce K. Howard, EN.

# ACKNOWLEDGEMENTS

Sincere appreciation is extended to Dr. Billy H. Johnson and to Dr. Victor L. Zitta for serving as major advisors during my thesis research and for their guidance, advice, and support. Sincere gratitude is also extended to Dr. Joseph H. Sherrard for serving as a member of my thesis committee and for his support in the successful completion of this research.

My appreciation is also extended to the U.S. Army Engineer Waterways Experiment Station for providing the opportunity to pursue this research, and particularly to Mr. William D. Martin whose support has been invaluable.

The author is also profoundly grateful to Dr. E. James Nelson, Dr. Norman L. Jones and the staff of the Engineering Computer Graphics Laboratory at Brigham Young University for their development of a watershed modeling software package called GeoShed with which the generation of TINs, delineation of watersheds, and computation of watershed geometric parameters was performed.

Finally, a heartfelt expression of appreciation is extended to my wife, Leah, and my sons, Will and Adam. This effort would not have been possible without their support, patience, encouragement, and understanding.

# CHAPTER I

# INTRODUCTION

Generating an accurate digital terrain model can be one of the most important aspects of computerized hydrologic modeling. From that terrain model, many characteristics of the area being modeled, such as the watershed boundary, surface area, and average slope, can be computed. There are several methods of constructing digital terrain models, and two of the major techniques are the grid-based method and the triangulated irregular network (TIN) method. With the grid-based method, elevation data points are spaced at regular intervals and the land surface is represented with rectangular shaped elements. In the triangular method, irregularly spaced elevation points are used to generate a continuous surface of interconnected triangles where the elevation data points serve as the vertices of those triangles. That network of interconnected triangles is referred to as a TIN. Each of these two methods has certain advantages and disadvantages.

Due to the fact that the data points are located at regular intervals, grid-based methods are very convenient from a computational and data storage viewpoint. However, the grid structure may not always provide the best basis

for a digital terrain model. The very structure of a grid dictates that elevation points in the grid will not always capture significant features in the terrain. The locations of peaks, ridges, valleys, and other important landscape features will most likely not conform to the regular pattern of the rectangular grid. To ensure that the smallest terrain feature is captured in a grid model, the density of data points in the entire grid must be increased to the level of detail required to define that feature (Peucker et al. 1976). That increased density of data points may result in gross redundancy of information in flatter areas of the land surface. In the application of terrain modeling for watershed analysis, there are additional problems with the grid structure. When upslope flow paths are computed, they follow the somewhat unrealistic zig-zag pattern of the grid network, and precision is often lacking in the definition of specific watershed boundaries (Moore, Grayson, and Ladson 1993).

The TIN method, however, overcomes many of these problems. The points which form a TIN are not required to conform to any particular structure. The points which form the TIN are ideally located where there is a change in slope, such that areas of constant slope are located on the planar surface of a triangle. Thus, in areas with rapidly changing terrain, a more dense sampling of points can be accommodated in a TIN while a more sparse density of points

in the same TIN can be used to represent flatter terrain. With appropriately selected points, a TIN can provide an accurate representation of the land surface with much fewer points than a grid would require. In a comparison of grids versus TINs, Peucker et al. (1976) found that a grid required more than 8 times the number of points than a TIN required to reflect the same degree of accuracy.

The particular application of TINs in this research is delineation of watershed boundaries and the computation of watershed geometric parameters for use in hydrologic modeling. For this application, TINs are very well suited and can perform better than a grid (Defloriani et al. 1986, Jones and Nelson 1992a). The TIN method does have some disadvantages as well. Since the points are not located in any regular or predictable pattern, the data storage structure is much more cumbersome than grid-based data for the same number of points. However, this is not necessarily a major disadvantage because much fewer points are required in a TIN. Also, the ideal location of points in a TIN is where there are changes in slope, but the availability of that type of data is limited, particularly when large watershed areas are being modeled.

For hydrologic modeling applications, the TIN structure provides the engineer a superior representation of the land surface with fewer points and provides an excellent framework for watershed analyses. However, for

engineers using TINs, the vast majority of elevation data currently available exists in a gridded structure. Thus, when a TIN model is desired, the problem arises as to how that gridded data may best be used to generate a TIN based terrain model from which input data for hydrologic modeling can be determined. The key to using gridded data for generation of a TIN is to identify those points in the elevation grid which most closely represent peaks, pits, valleys, ridges, and other significant changes in slope. Once those points have been identified, they can be used to serve as the vertices for a TIN model. The remaining points in the grid are generally redundant and their inclusion in the TIN does not appreciably increase the accuracy of the TIN model, thus those remaining grid points may be discarded.

There are several methods for selecting appropriate points in gridded elevation data from which to generate a TIN model. Each of these methods effectively filters out a certain number of the elevation points in the grid that are determined to be "insignificant" based upon some user defined threshold. Typically, the user has control over how much data is filtered out of the grid before the TIN is generated, and it is up to the user to determine how much data filtering is appropriate for a given situation. Providing guidance in the appropriate use of one particular filtering method is the focus of this research.

The particular filtering method analyzed here is a curvature based technique, the details of which are contained in Chapter V of this report. This curvature based technique allows the user some control over how many points in the elevation grid are selected to generate a TIN. This can result in a TIN using as few as approximately 5% or as many as perhaps 80% of the original grid data points. Obviously, the resultant TIN will be different if it is generated with only 5% of the original data points than it would be if it were generated with 80% of the original points. The question to be answered by this research is how those TINs actually do differ as less data points are used, and what are the impacts of this data filtering when the ultimate goal is to generate input for use in a hydrologic model.

To answer this question, the following steps were followed in this research and are the subject of the remainder of this report:

1. Gridded elevation data were obtained for three watershed areas, each with a different type of topography.
2. The gridded data were filtered using the curvature based technique mentioned above, and the filtering was performed at several thresholds for each set of data such that various amounts of elevation data were selected for TIN generation.

3. TIN models were generated from each set of filtered data.
4. Those TINs were analyzed to determine the elevation differences, or error, as compared to the original gridded data. Statistical analysis was performed on the error data to help measure the quality of each TIN surface.
5. Using the TINs generated in Step #3, watershed boundaries were delineated and various geometric parameters for those watersheds were computed from the TIN geometry. The differences in those geometric parameters versus the amount of data filtered from the grid were determined.
6. The geometric parameters computed in step #5 were used as input for the HEC-1 hydrologic model, and the differences in the resultant hydrographs versus the amount of data filtered from the grid were documented.

By following the above steps, the effectiveness of the data filtering technique selected was analyzed, and guidance to potential users on the appropriate level of data filtering for use in hydrologic modeling applications provided. The user may refer to this work to determine the impact of filtering data at a particular threshold, and to seek guidance on the level of filtering which is appropriate for a given type of topography. The research

process outlined above is detailed in the following Chapters after some preliminary information and discussion of hydrologic modeling and TINs.

# CHAPTER II

# BACKGROUND

For centuries, engineers have attempted to predict the amount of flow that will appear in a stream or river as the result of a particular storm. In very broad terms, this is the basic issue addressed by the science of hydrology. It has been said that the science of hydrology began with the concept of the hydrologic cycle, which is shown conceptually in Figure 1. Although the earliest versions of the hydrologic cycle were primitive and scientifically unsound, a reasonably accurate concept of the hydrologic cycle was formulated by the Roman architect and engineer Marcus Vitruvius who lived about the time of Christ. The concept of the hydrologic cycle may also have been developed independently in China by 900 B.C. and in India by 400 B.C. (Chow, Maidment, and Mays 1988).

Although these basic hydrologic concepts have been around for centuries, quantitative hydrology has undergone most of its development during the twentieth century and hydrology has been recognized as a separate discipline only in relatively recent years. Perhaps the most revolutionary impact on hydrology has been the advent of the computer age. Theories and methods that were virtually impossible

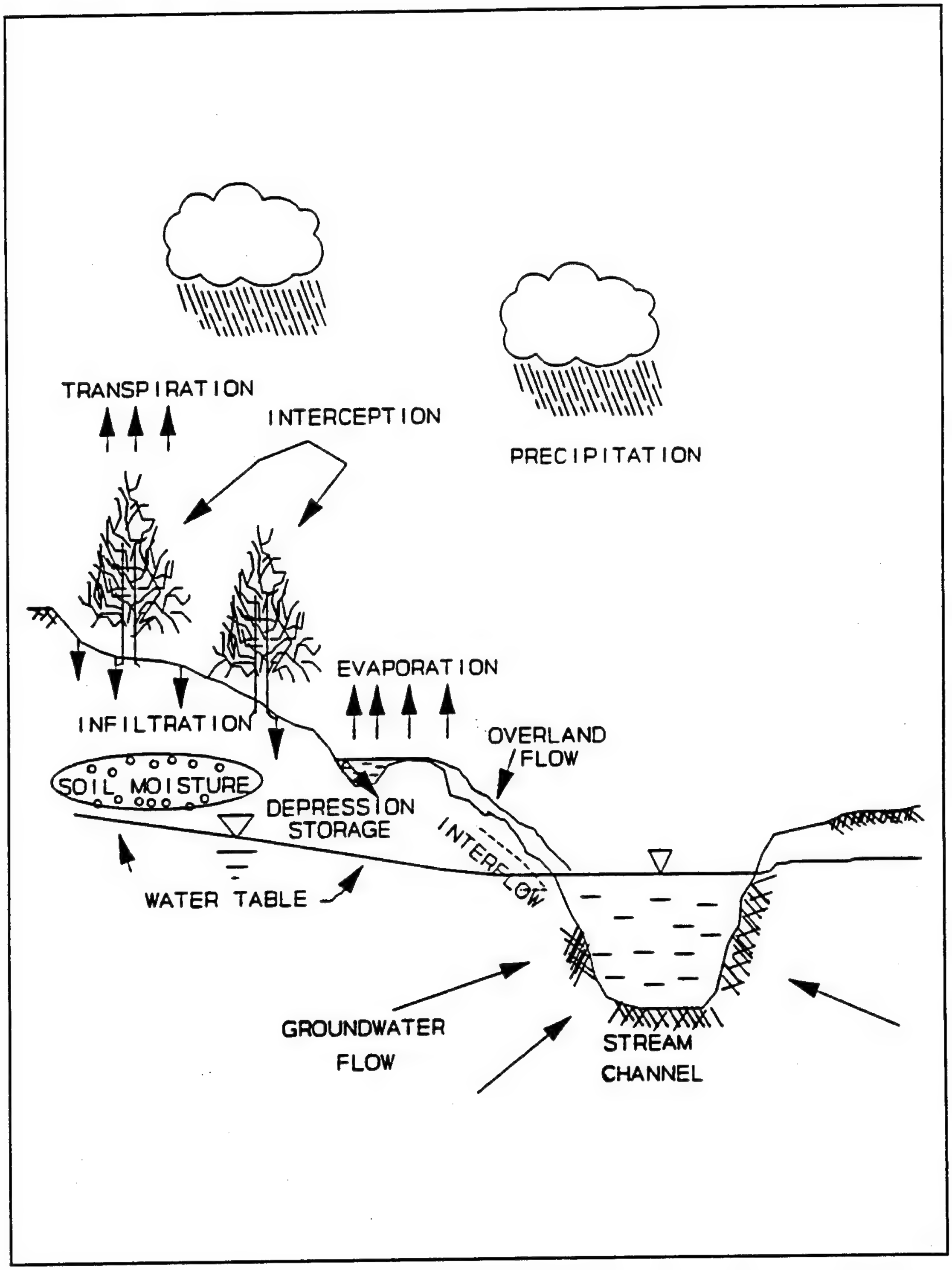



Figure 1. The Hydrologic Cycle (Department of the Army 1993)

to apply to hydrologic problems in earlier times are now routinely applied through the speed of modern computers, and a host of computerized hydrologic models have been developed to take advantage of this technology. Hydrologists are now able to deal with problems on a much larger scale using larger amounts of data than have ever been possible before. However, with that increased capability comes new issues to be addressed. For example, issues of scale, data resolution, digital data accuracy and formats, and terrain modeling techniques in the field of hydrology today were not issues at all until relatively recent times.

## Hydrologic Modeling

Prior to the advent of the computer age, a typical hydrology problem was completed using relatively simplistic techniques. However, with the aid of computers, hydrologic analysis is now almost always performed through the use of computerized numerical models. Although the use of computerized hydrologic models has become commonplace, it is a field that is continuously changing as computer technology changes.

Hydrologic modeling is currently undergoing tremendous changes, not only in the development and application of new computational techniques to compute runoff but in the development of new tools which enable more efficient use of existing techniques. These changes are due to many

factors, including the advancement of computer hardware, the development of Geographic Information Systems (GIS) for data management, and the availability of data with which to conduct hydrologic modeling.

## Types of Models

There are many different ways in which hydrologic models can be classified. Two of the general classes of models applied frequently in hydrology today are lumped parameter models and distributed models (Chow, Maidment, and Mays 1988). Lumped parameter models use average values for hydrologic parameters over relatively large areas, typically subbasins in a watershed, while distributed models use much smaller computational elements which more fully define the true spatial variability of the system.

Historically, distributed models have not been as widely used as lumped parameter models, mainly due to the vast computational requirements of distributed modeling. Although distributed models are becoming more feasible as computers become faster, lumped models remain probably the most common type of model in general use today. Many of these lumped models have been in use for several decades and are well proven. Thus, the enhancements to these existing models that are arising from current technology are not generally due to new theoretical developments or computational methods but in the use of new data sources and GIS systems to manage the data.

## Model Input

The input parameters for lumped models have historically been developed manually from topographic maps, soil surveys, and field reconnaissance. However, much of the data necessary for running the models are now available in digital form. Specifically, the United States Geological Survey (USGS) distributes elevation, land use / land cover, and stream location data in digital formats, and the Natural Resources Conservation Service (NRCS), formerly the Soil Conservation Service (SCS), is in the process of making digital soil data available throughout the country. In addition, there are many other federal, state, local, and private organizations through which data are available. An important issue in the field of hydrologic modeling that is addressed in this research is how to make the best use of these newly available digital data sources in existing models.

This research focuses specifically on the use of newly available gridded elevation data in the existing HEC-1 hydrologic model. However, the general conclusions of this research are not limited to HEC-1, but instead are applicable to any lumped parameter model. Thus, this research provides guidance to assist hydrologists in making the best use of available digital elevation data to create TINs and generate a sound engineering product in an efficient manner.

## CHAPTER III

## THE TRIANGULATED IRREGULAR NETWORK

This research involves the triangulated irregular network (TIN) method of terrain modeling, so a detailed discussion of TIN models is necessary and appropriate. To construct a TIN, data points with x,y,z coordinates are required. The x,y,z values, typically in a rectangular coordinate system, represent the easting (x), northing (y), and elevation (z) of a point. The points are connected to form a network of triangles where the points serve as the vertices of the triangles. Figure 2 depicts a sample TIN constructed from a set of scattered data points. Once the

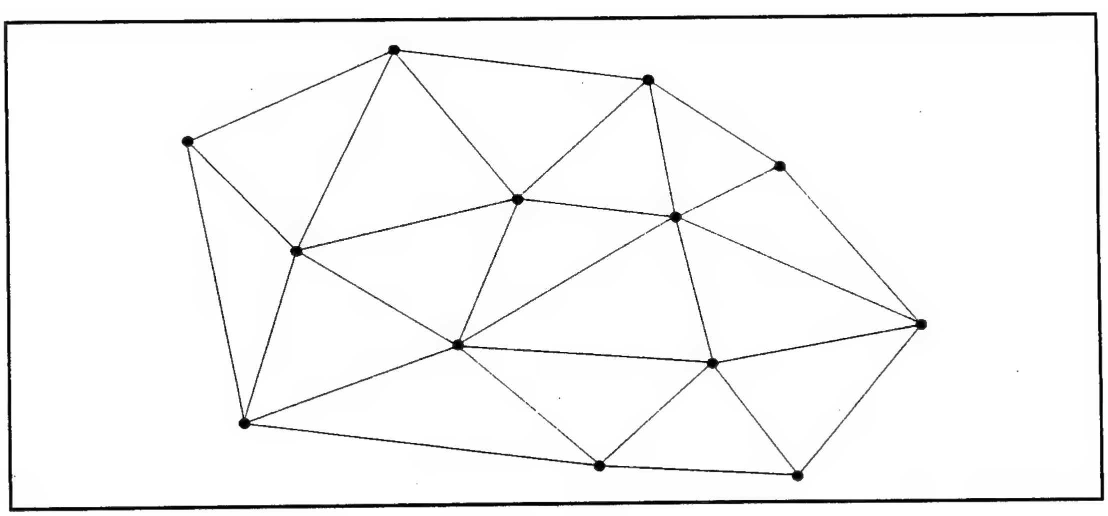


Figure 2. Sample TIN

triangles have been formed, the elevation at any point within the TIN is determined by linear interpolation across the face or edge of the triangle within which that point is located. Thus, for any given x,y location, the corresponding z (elevation) is easily determined.

## Triangulation

Given a set of points, there is more than one way in which a triangular network can be formed between those points. For example, Figure 3 includes four points and shows that there are two different ways in which those points can be connected to form a triangular network since the diagonal line connecting the points can be oriented in two different ways. Thus, it is desirable to have a standard method for generating a TIN such that

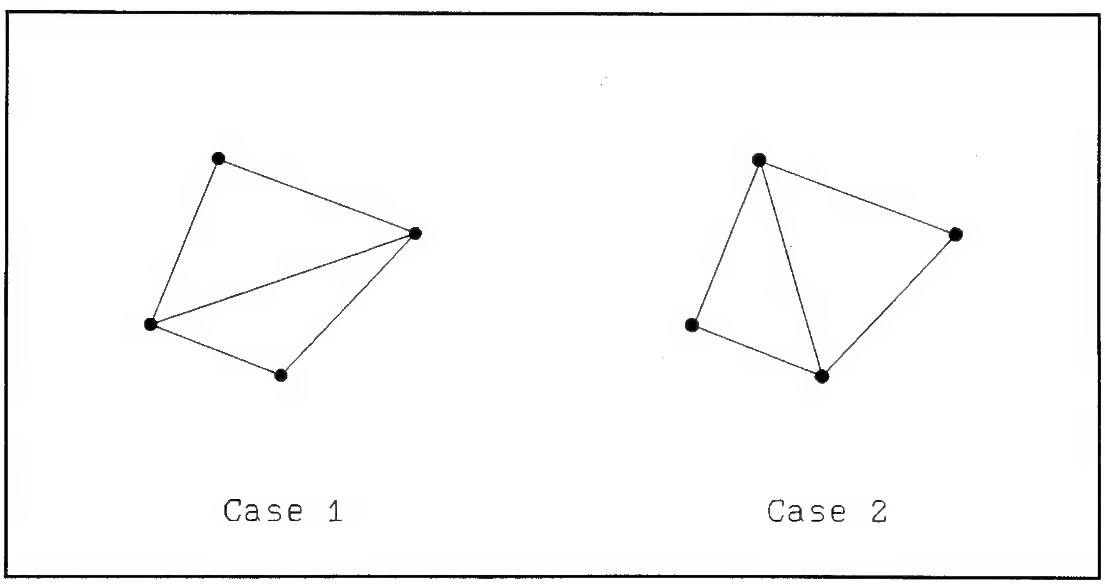



Figure 3. Two Possible Triangle Configurations

triangulation of any given set of data points will always result in the same triangular configuration or TIN.

Many different algorithms have been presented for triangulating a set of points to form a TIN (Lawson 1977; Lee and Schacter 1980; Watson 1981; Watson and Philip 1984; and Jones 1990), and most of those methods use the Delauney criterion to guide the triangulation process. The Delauney criterion is satisfied when the circumcircle of the three vertices of a triangle does not encompass any other vertices, where the circumcircle is the circle defined by the three corners of the triangle as shown in Figure 4.

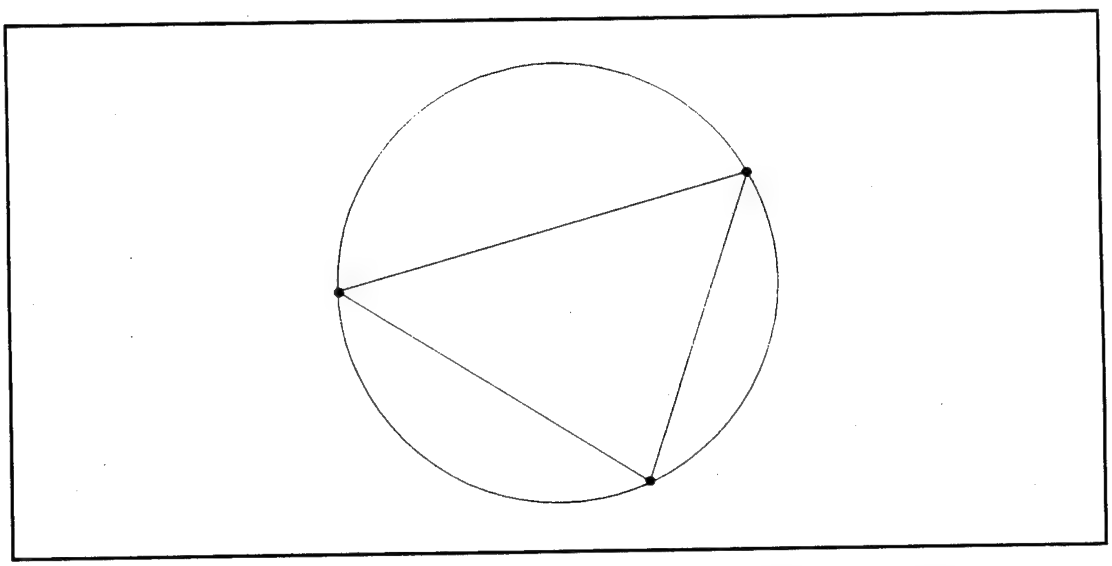


Figure 4. Triangle and its Corresponding Circumcircle

Figure 5 is used to illustrate the Delauney criterion. Note that in Case 1 of Figure 5, the circumcircles both encompass not only the three vertices that define their

triangle, but they also encompass the other vertex. Thus, the Delauney criterion is violated. In Case 2 of Figure 5, both of the circumcircles encompass only the three vertices which define their triangles, and the Delauney criterion is therefore satisfied. So in this example, Case 2 would be the appropriate configuration for a Delauney triangulation. The end result of creating a TIN using the Delauney criterion is that the existence of long, thin triangles is minimized and the triangles that do make up the TIN are as equiangular as possible (Nelson, Jones, and Miller 1994).

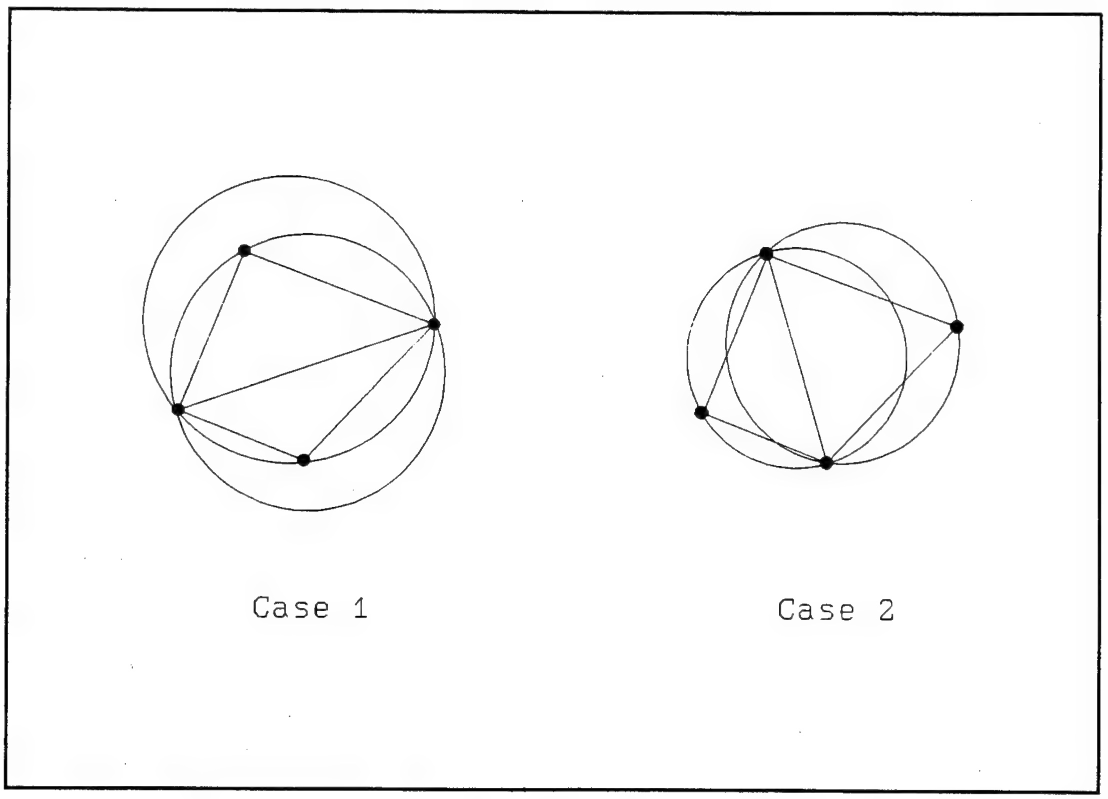



Figure 5. Circumcircles for Two Triangle Configurations

## TIN Data Format

As noted previously, there are many basic differences between grid-based terrain models and TIN-based models. One of those differences is obviously the format of the data file. Gridded data can be stored in a data file much more efficiently than can data in a TIN. Gridded elevation data files generally contain information on the number of rows and columns of data and the spacing between those rows and columns followed by a listing of the elevation data values. TINs, on the other hand, have no such regular spacing between points and the data is not organized into any rows or columns. Rather, the TIN is constructed from irregularly spaced points, and the x,y location of each point must be specified along with the elevation, or z, value at that location.

In some of the discussions related to the uses of TINs later in this document, an understanding of the data format of the TIN file will be necessary and is introduced here. The basic TIN data file format used in this research is broken down into two major sections. The first section is vertex data containing the x,y,z information for all of the points which serve as vertices of the triangles in the TIN. The second section includes triangle data which define the three points that are connected to form each triangle in the TIN. To illustrate this TIN file format, a sample data file in this format is shown below. In this sample data

file, NUMV represents the number of data points in the file, $x_1, y_1, z_1$ are the x,y,z coordinates of the first point in the file, NUMT is the number of triangles, and $v_{11}, v_{12}, v_{13}$ are the three vertices which form the first triangle:

VERT NUMV
$x_1,\ y_1,\ z_1$
$x_2,\ y_2,\ z_2$
$x_3,\ y_3,\ z_3$
.
.
.
$x_{NUMV},\ y_{NUMV},\ z_{NUMV}$
TRI NUMT
$v_{11},\ v_{12},\ v_{13}$
$v_{21},\ v_{22},\ v_{23}$
$v_{31},\ v_{32},\ v_{33}$
.
.
.
$v_{NUMT1},\ v_{NUMT2},\ v_{NUMT3}$

To further illustrate this format, Figure 6 shows a sample TIN which was generated by Delauney triangulation of five elevation data points, or vertices. For reference purposes, each point is labeled with a vertex number and elevation, and each triangle is numbered. Following the TIN data file format shown above, the TIN file for the data shown in Figure 6 would be as follows:

```
VERT 5
2, 3, 88   (note: x=2, y=3, z=88)
1, 9, 90
4, 5, 91
7, 7, 98
7, 2, 95
TRI 4
1, 3, 2    (note: vertices #1, #3, #2 form triangle #1)
2, 3, 4
3, 5, 4
1, 5, 3
```

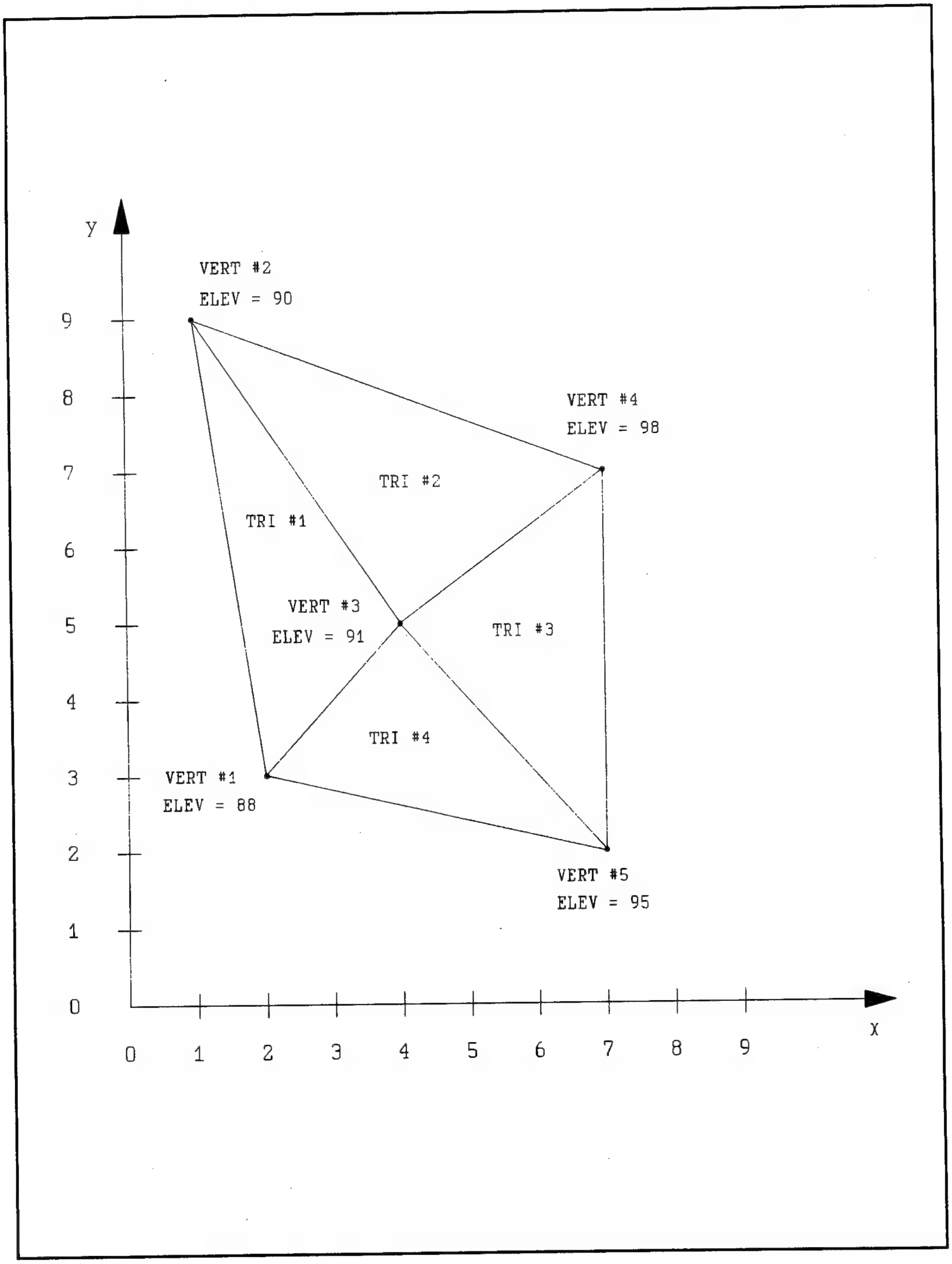



Figure 6. Sample TIN to Illustrate Data Format

Notice that the vertices which form each triangle are stored in counterclockwise order in the TIN file. This provides a consistent format and simplifies many of the computations that are performed with the TIN data. If there were no particular order to the vertices in the triangle data of the TIN file, then some of those computations would not be possible.

## Watershed Analysis Using TINs

Once a TIN has been generated using the methods outlined above, the hydrologist often wishes to use that TIN to help in the process of watershed analysis. Methods have been developed which use the geometric information contained in the TIN to define the areas which contribute flow to a given point, and these methods can be used to easily delineate watershed boundaries. After the watershed is delineated, the TIN serves as an excellent basis for computing geometric properties of the watershed such as area, average slope, and flow distances.

### Watershed Delineation

Jones, Wright, and Maidment (1990) proposed an algorithm for automated watershed delineation using TINs. While that technique worked well in most cases, there were several shortcomings. Nelson, Jones, and Miller (1994) addressed those shortcomings and set forth a new algorithm that precisely delineates watershed boundaries using a TIN

model. This algorithm is used to define the boundary of an area contributing to the flow at a single point or for multiple watersheds in a stream network. This algorithm, as set forth by Nelson, Jones, and Miller (1994), is the one used in this research and is detailed below.

The most fundamental aspect of the watershed delineation method is tracing flow paths on the TIN. Assuming that roughness and momentum are negligible, the direction of flow of water across a surface will be in the direction of steepest descent, i.e., the direction of the maximum downhill gradient. Jones, Wright, and Maidment (1990) described this process and showed that flow paths can be constructed from any arbitrary point on a TIN by following the path of maximum downward gradient from triangle to triangle. The path of flow is orthogonal to the contour lines on any given triangle.

This is illustrated in Figure 7 which shows several triangles with contour lines, and a sample flow path along the path of maximum downward gradient. Notice that in Figure 7 all of the flow occurs across the faces of the triangles. If the flowpath were to have intersected an edge between two adjacent triangles which both slope towards each other, then the flow would have continued downward along that edge. By following this succession of flow, either down across the triangle faces or down along triangle edges and always along the path of steepest

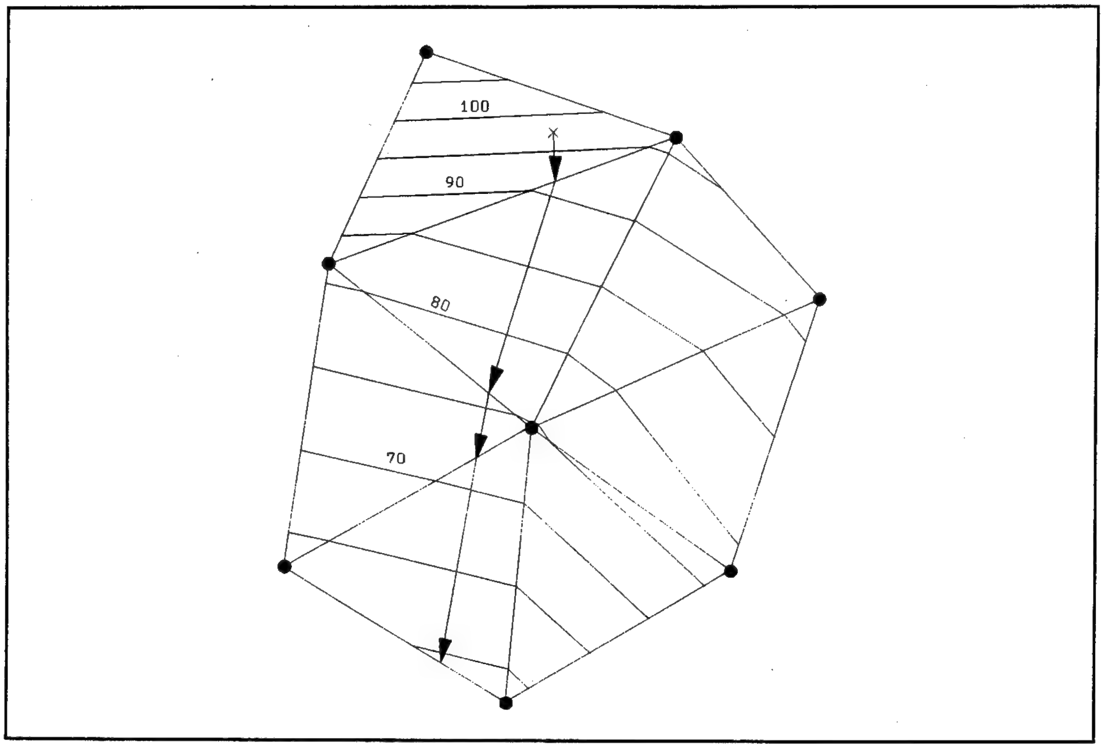



Figure 7. Path of Maximum Downward Gradient

downward gradient, the flowpath can be traced to a terminus. That terminus can be either a pit or local depression in the interior of the TIN, a boundary on the exterior of the TIN, or a user defined point at any location in the TIN.

When a flowpath is initiated at the centroid of each triangle in the TIN, the drainage patterns in the entire TIN become evident. Figure 8 shows a sample TIN which was triangulated using the Delauney criterion from a set of scattered x,y,z data points and the contours on that TIN are shown in Figure 9. Using that same TIN, the drainage pattern which resulted from initiation of a flow path at the centroid of each triangle is depicted in Figure 10.

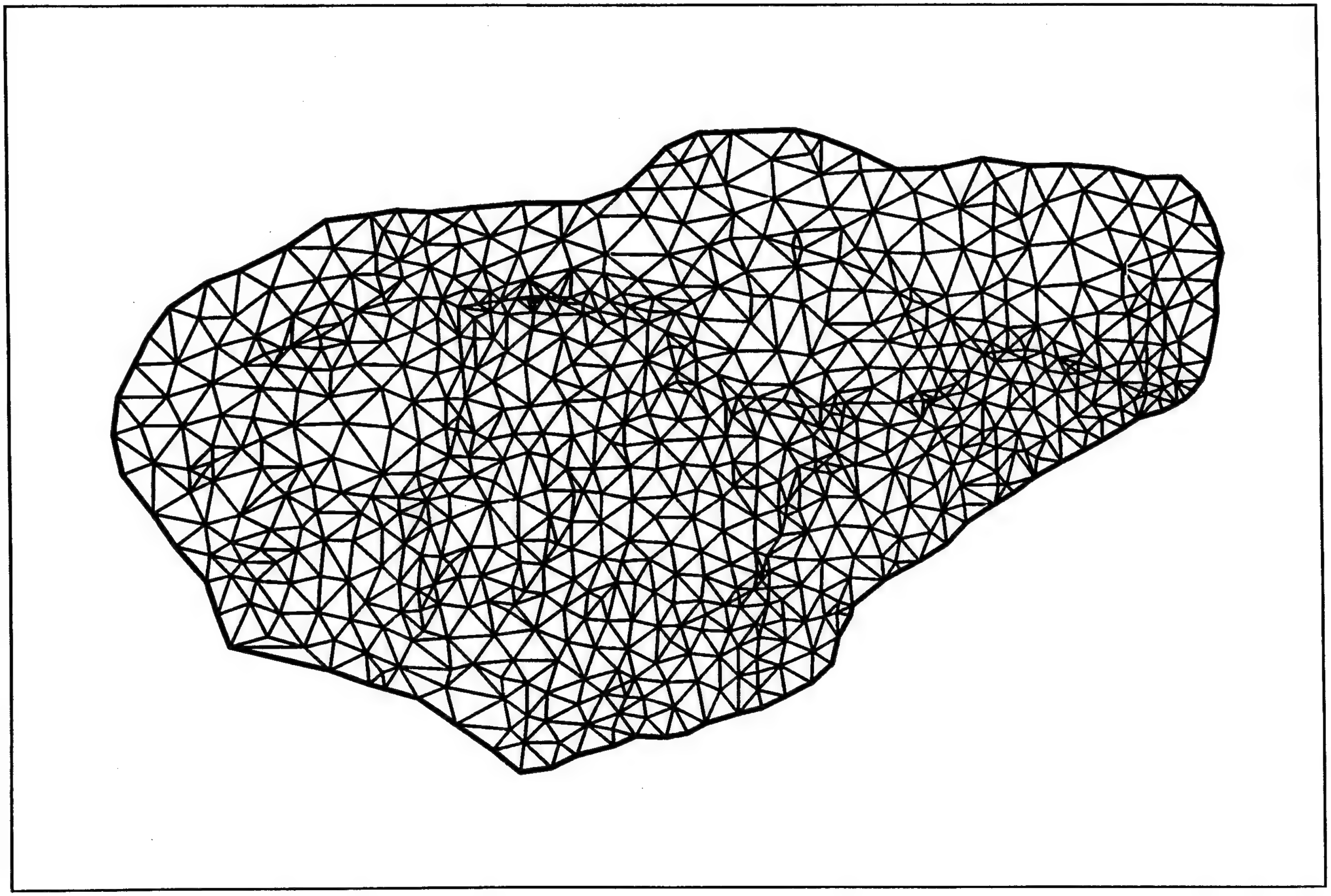


Figure 8. TIN for a Sample Watershed

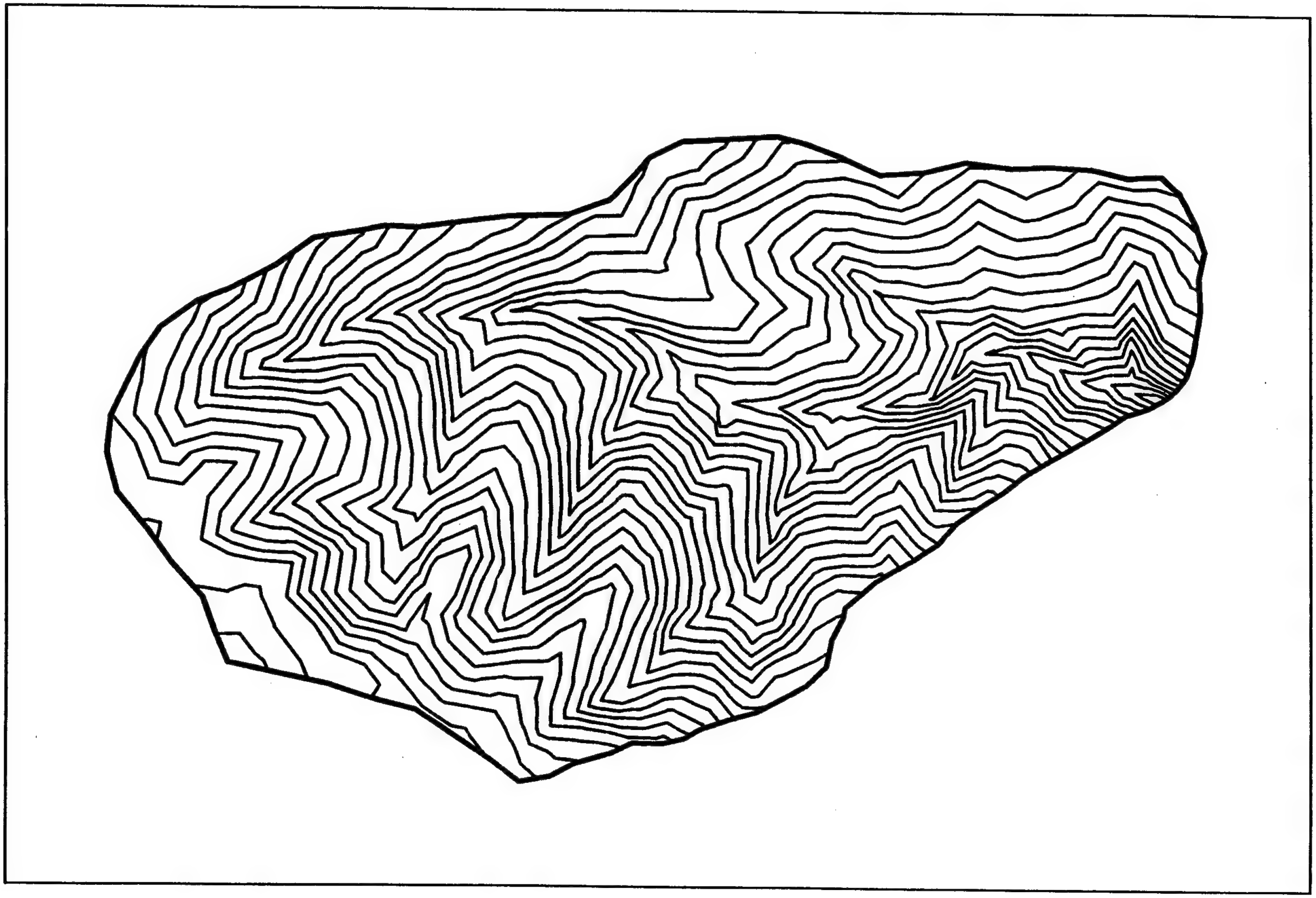


Figure 9. Contours for Sample TIN

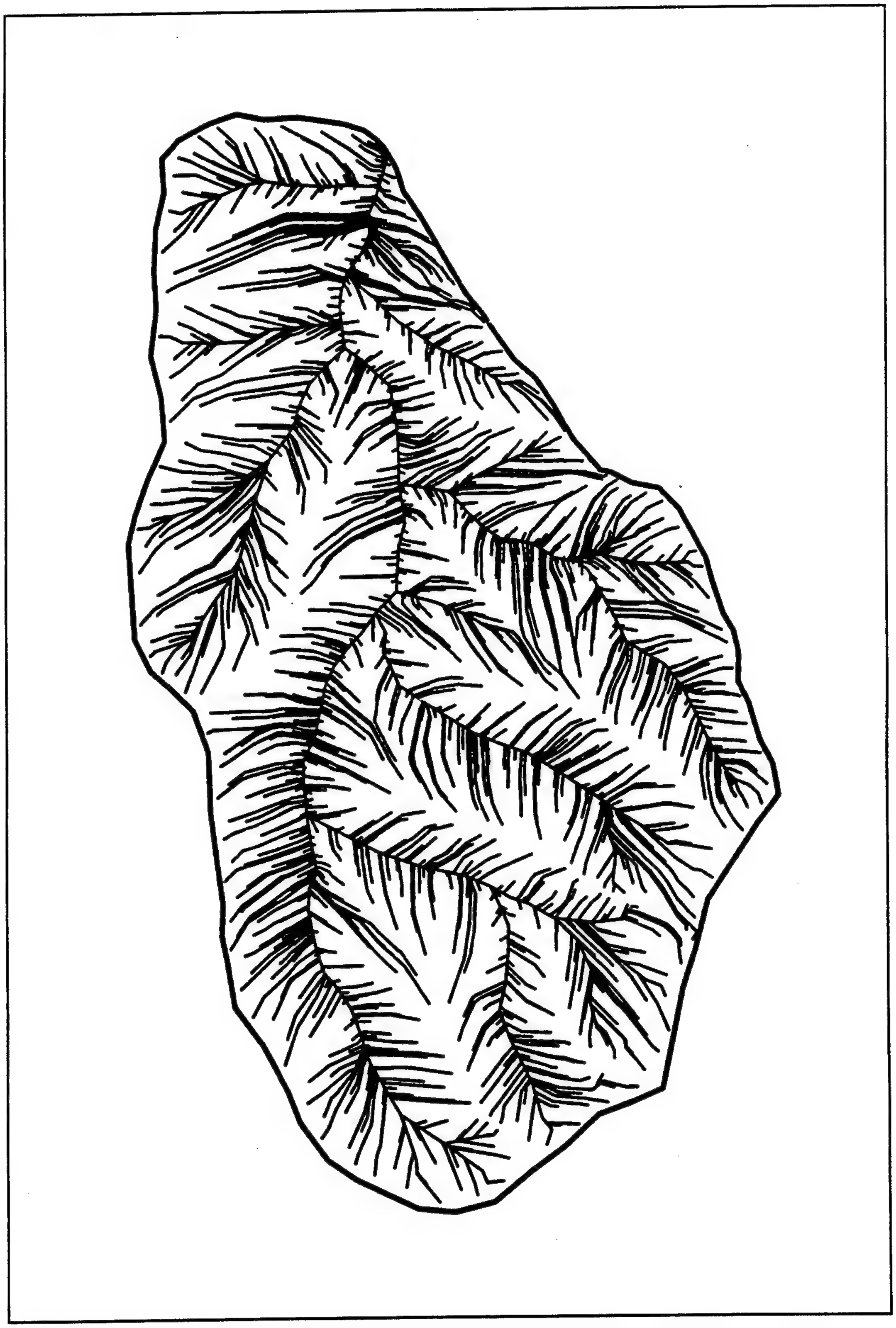


Figure 10. Flow Paths on Sample TIN

In this example, the terminus is the outlet of the basin, but a terminus could also be defined by the user as any point in the TIN. When a flowpath intersects a terminus point, that particular flowpath terminates at that point.

The drainage basin or watershed area for any given terminus point is determined by defining the set of triangles that contribute flow to that terminus. This set of triangles is determined by generating a flowpath at the centroid of each triangle in the TIN and following each flowpath in the direction of maximum downward gradient until it intersects a terminus. All such triangles whose flowpaths intersect a given terminus are then said to contribute flow to that terminus. This is then the set of triangles which comprise the drainage basin for that terminus, and the perimeter of that set of triangles is defined as the watershed boundary. In most cases, there are triangles whose flow intersects the TIN boundary before it reaches a terminus point. Those triangles are not considered to contribute to any basin, and may either be ignored or discarded from the TIN.

Figure 11 shows the same TIN as presented in Figures 8, 9, and 10, but with several terminus points within the TIN identified. One terminus is actually located at the outlet of the basin and is on the boundary of the TIN. The two terminus points located in the interior of the basin have been specified as outlet points for subbasins. As can

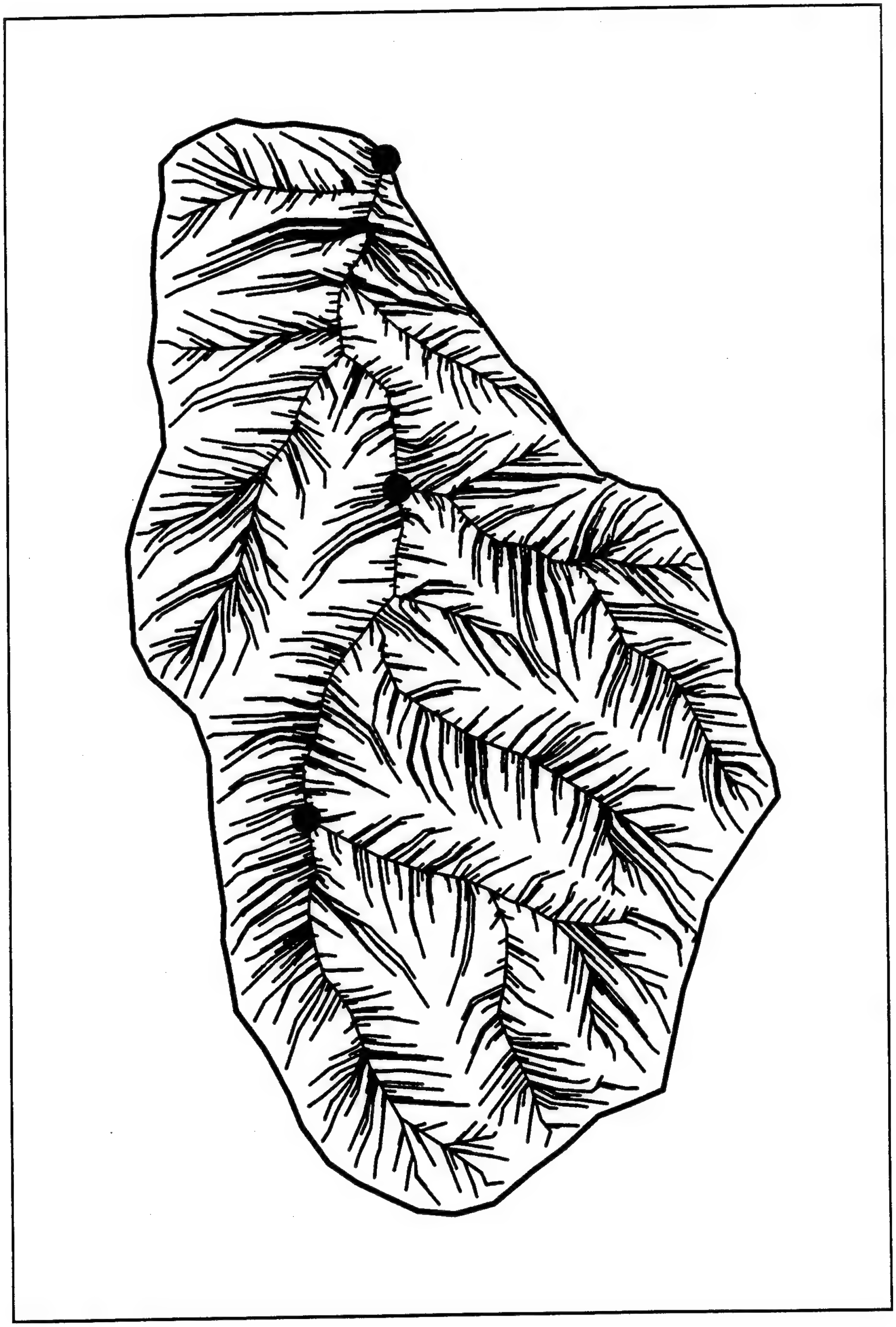


Figure 11. Terminus Points within Sample TIN

be seen by studying the drainage pattern in the TIN, those two termini are located at branching points in the stream network. Using those specified terminus points, the drainage area associated with each terminus was identified as the set of triangles which contributed flow to that point. The resultant drainage areas are shown as five subbasins in Figure 12. Using this technique, the user can specify any number of terminus points and have the drainage area for those points quickly identified.

## Basin Geometric Parameters

After defining a watershed area as described above, the underlying TIN serves as an excellent base from which to automatically compute the geometric parameters for that watershed required in a hydrologic model (Nelson, Jones, and Miller 1994). The TIN is comprised entirely of a set of triangles, and each of the triangles in the TIN is defined by three vertices with known coordinates. Thus, the information required to compute geometric properties such as area, slope, flow distances, and centroid is contained in the TIN, and the computation of those geometric properties is a fairly straightforward procedure. In this research, three specific geometric properties were considered for each watershed. Those properties are the area, the average slope, and the maximum flow distance within the basin. A discussion of each of these geometric properties follows.

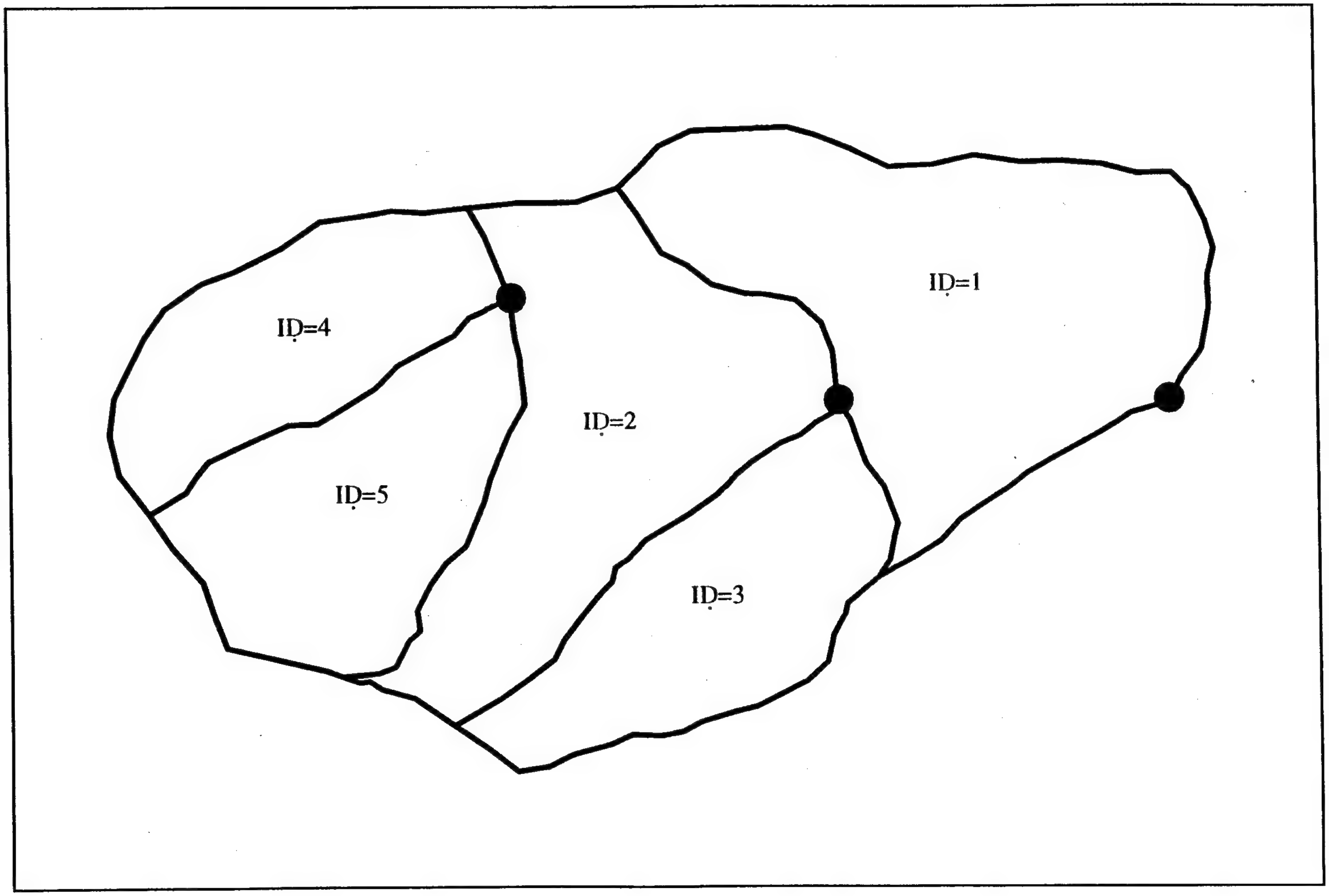



Figure 12. Drainage Areas for Selected Terminus Points

Area

The area of any given triangle is computed with the following general equation for the area of a triangle

$$Area = \frac{1}{2}\, a\, b\, (\sin C) \tag{1}$$

where a, b, and C are defined in Figure 13 and are easily determined from the coordinates of the points in a TIN. The area for a watershed is determined by simply summing the areas of all triangles which belong to that basin.

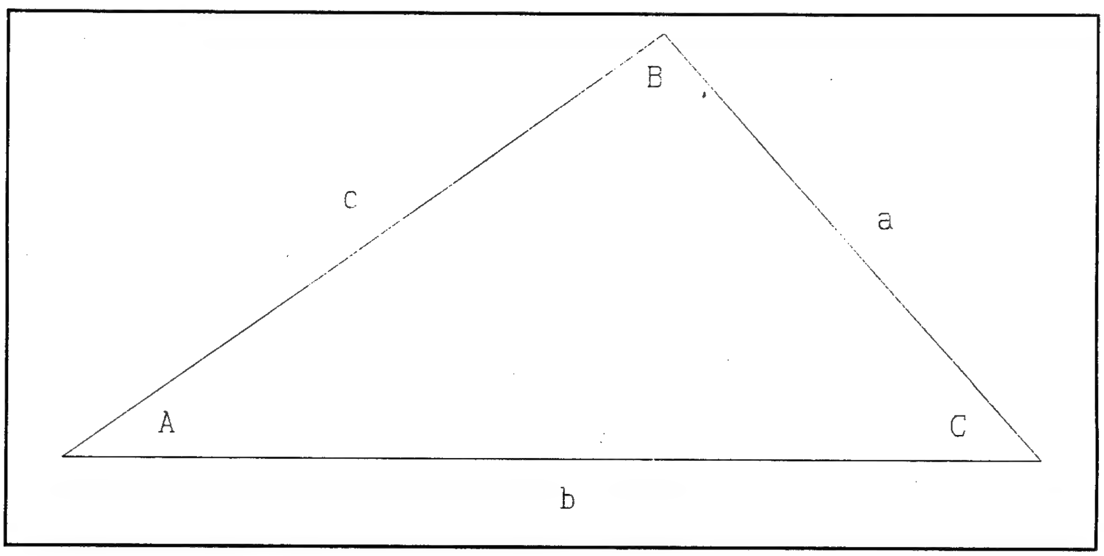



Figure 13. General Triangle

Average Slope

The average slope of any given triangle in the TIN is determined by calculating the slope of the plane in which that triangle is located, where the plane is defined by the three vertices of the triangle. Take, for example, the sample triangle shown in Figure 14. The three points which

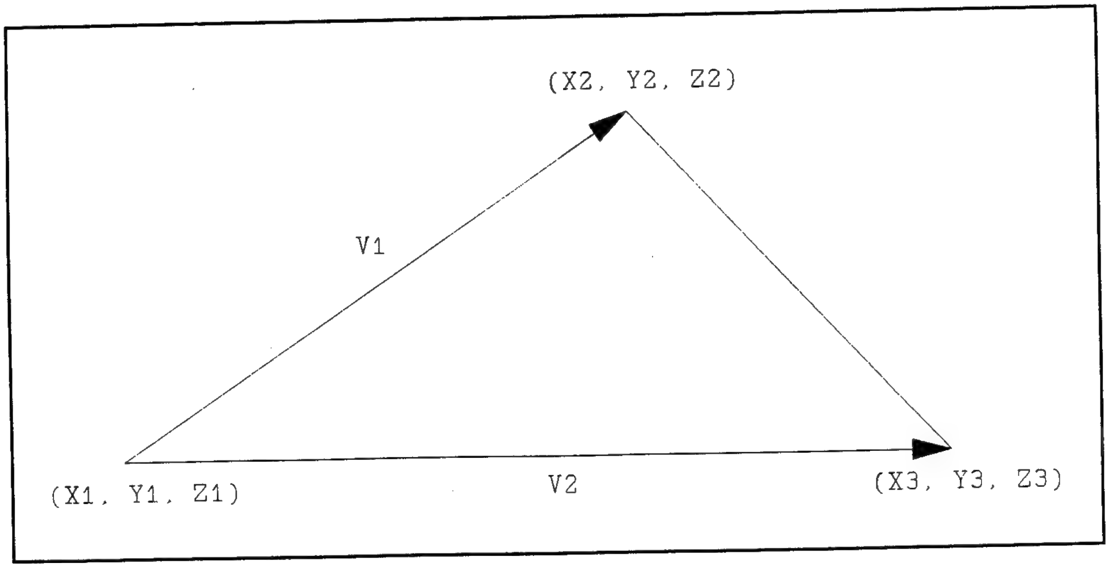



Figure 14. Vector Definition

serve as the vertices of this sample triangle have coordinates (X1,Y1,Z1), (X2,Y2,Z2), and (X3,Y3,Z3). The vector V1 is defined as the vector originating at (X1,Y1,Z1) and leading to (X2,Y2,Z2), vector V2 is the vector originating at (X1,Y1,Z1) and leading to point (X3,Y3,Z3), and these can be written as:

$$\vec{V1} = (X2 - X1)\hat{i} + (Y2 - Y1)\hat{j} + (Z2 - Z1)\hat{k} \tag{2}$$

$$\vec{V2} = (X3 - X1)\hat{i} + (Y3 - Y1)\hat{j} + (Z3 - Z1)\hat{k} \tag{3}$$

The slope of the triangle is determined by calculating the angle between the plane containing the triangle and the horizontal plane. Determining the slope of the plane relies on the fact that the angle between any two planes is the same as the angle between their normal vectors. The

normal vector, $N_{TRI}$, for the plane containing the triangle can be determined by taking the cross product of vectors V1 and V2, i.e.,

$$\vec{N}_{TRI} = \vec{V1} \ X \ \vec{V2} \tag{4}$$

The normal vector to the horizontal, $N_{HOR}$, can be defined as:

$$\vec{N}_{HOR} = 0\hat{\imath} + 0\hat{\jmath} + 1\hat{k} \tag{5}$$

Given these two normal vectors, the angle, $\phi$, between the two planes is determined by using the following equation:

$$\cos \phi = \frac{\vec{N}_{TRI} \cdot \vec{N}_{HOR}}{|\vec{N}_{TRI}| \ |\vec{N}_{HOR}|} \tag{6}$$

Taking the $\cos^{-1}$ of this value then gives the angle, $\phi$, between the triangle and the horizontal plane. Now, the slope of the triangle is simply:

$$SLOPE = \tan \phi \tag{7}$$

By following this process, the slope for each triangle in the watershed is computed. The average slope for the watershed is determined by computing a weighted average of the slopes based upon the relative area of each triangle in the basin.

Maximum Flow Distance

Earlier in this Chapter, the method of determining the flowpath on a TIN from a given point was outlined. That method involved following the path of maximum downward gradient from a given point to a terminus. The maximum flow distance in a watershed is taken by comparing the lengths of all flowpaths initiated from the centroid of each triangle in the basin to the terminus. The longest of those is taken as the maximum flow distance for the basin.

## TIN Summary

Based upon the methods and techniques presented in this Chapter, TINs can effectively and efficiently be used to generate terrain models from which watershed boundaries and geometric parameters can be estimated. Figure 15 shows the basin boundaries, areas, average slopes, and maximum flow distances for subbasins of the watershed shown in Figure 12. For the research presented herein, a curvature based filtering technique was used to select points from gridded elevation data to generate TINs. Then, the methods outlined in this Chapter were used on those TINs to perform watershed delineation and to compute basin geometric parameters for use in the HEC-1 hydrologic model. Discussions of the type of gridded elevation data generally available, details of filtering of those data, and creation of TINs from the filtered data are presented in the following chapters.

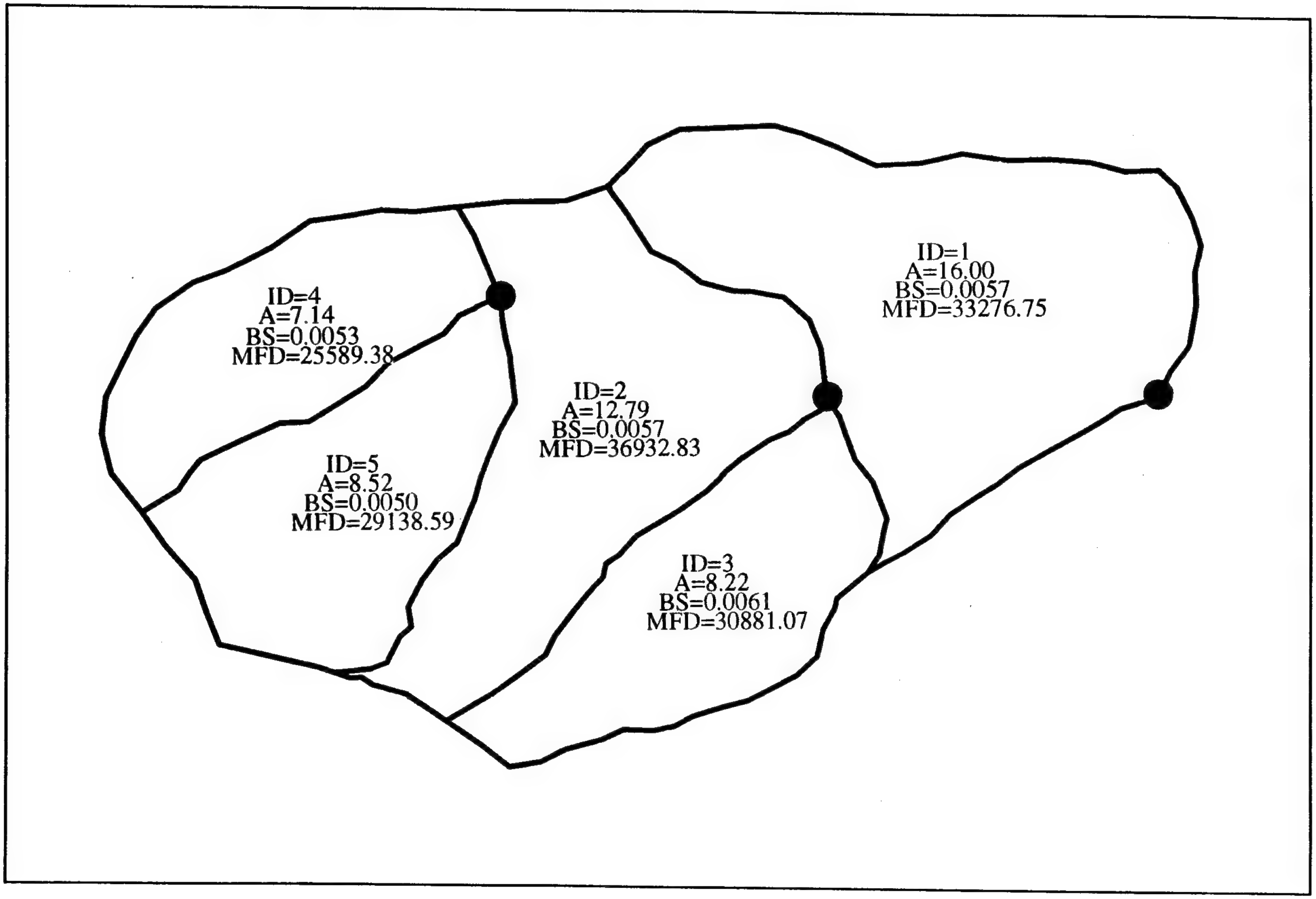



Figure 15. Geometric Parameters for Subbasins

# CHAPTER IV

# ELEVATION DATA

Perhaps the single most important type of data that is used as input to any hydrologic model is the elevation data which represents the topography of the area being modeled. Not only is topography important in defining the movement of water in a system, it influences many aspects of the hydrologic system such as flow paths for precipitation, the spatial distribution of soil moisture, and the chemical characteristics of streamflow (Wolock and Price 1994). While this research does not address all of these issues, it is helpful to recognize that the elevation data, with which the topography in a watershed is represented, are extremely influential and important.

## Elevation Data in Hydrologic Models

Raw elevation data are typically not used directly in lumped parameter hydrologic models, but the elevation data are used to derive other information required as input to the model. Specific examples of such information are the watershed and subbasin boundaries and areas, slopes, and flow distances as discussed in the previous chapter. While there are many other very important variables in the

modeling process, accurate depiction of the size, shape, and slope of the watershed are critical.

From the discussion of TINs previously presented, it is obvious that the key to generating a TIN which accurately depicts the land surface is to have elevation data points appropriately located. The most advantageous location of elevation points for a TIN is to have a point at each location where a change in slope occurs. In this case, areas of constant slope are accurately represented by the surfaces of the triangles while the breaks in slope occur along the triangle edges. Realistically, this situation will almost never occur, particularly when modeling a natural watershed where true areas of constant slope are rare. Additionally, the task of acquiring an elevation data point at every location where a change in slope occurs would be virtually impossible in most situations. Thus, hydrologists must settle for some reasonable approximation of the terrain surface in the TIN model, and providing guidance on how good that approximation must be is the major goal of this research.

## Elevation Data for TINs

As stated earlier, a TIN can be generated from any set of x,y,z data points. The points may be located on a regular grid structure or may be irregularly spaced. Using a full set of gridded points with none of the data filtered out essentially eliminates many of the advantages of using

a TIN, since a TIN can represent a surface with a fraction of the points contained in a grid (Peucker et al. 1976).

The most desirable situation is to generate a TIN from irregularly spaced points. The sources of those irregularly spaced points may include data from a field survey, data obtained by digitizing a topographic contour map, or by choosing selected points from gridded data. Provided that the data are all maintained in the same coordinate system and properly georeferenced, any combination of the above data sources can also be used.

Typically, only one of the above data sources will be used in any given study. The most accurate of these would probably be a field survey, but this is a relatively expensive option. Obtaining detailed field survey data for large areas to be modeled is typically not a practical option. Digitizing contours from a topographic map is not necessarily expensive, but it can be tedious and is prone to injecting human error in the digitizing process. Also, the digitized points must be spaced properly along contours in order to generate the best TIN, and this process can be difficult to accurately perform. The type of elevation data that is becoming more prevalent are gridded data, which is probably the most common format for digital elevation data over large areas of the United States. These gridded data obviously serve the purpose of those who use grid based methods for modeling, and as shown in this

research, the gridded data can also be effectively used in TIN generation.

## Digital Elevation Models (DEMs)

The focus of this research is to provide guidance on the use of gridded digital elevation data for generating TIN models and the subsequent computation of watershed characteristics from those terrain models for use in hydrologic modeling. The most readily available source of gridded data for large areas within the United States is the DEMs from the United States Geological Survey (USGS). The USGS distributes two separate digital elevation products, namely, the 7.5 Minute DEM and the 1 Degree DEM. The 1 Degree DEM product was used for the research presented herein. Thus, only a brief description of the 7.5 Minute DEM product is provided while more detail is provided for the 1 Degree DEM data.

### 7.5 Minute DEM

The 7.5 Minute DEMs provide coverage for 7.5 x 7.5 minute areas corresponding to standard USGS 1:24,000 scale maps. A 7.5 Minute DEM consists of a regular array of elevations referenced to the Universal Transverse Mercator (UTM) coordinate system horizontally. Elevation data points are spaced on a 30 x 30 meter square grid. These DEM data are available for most areas within the United States and may be purchased from the USGS (USGS 1987).

1 Degree DEM

Each 1 Degree DEM covers a 1 x 1 degree area. The USGS 1:250,000 scale maps provide 1 x 2 degree coverage, and each DEM corresponds to the East or West portion of one of those maps. Figure 16 shows a sample 1 x 2 degree coverage for a 1:250,000 scale map and the corresponding 1 x 1 degree West and East DEMs for the same area. The basic elevation model for the 1 Degree DEM is produced by the Defense Mapping Agency (DMA), but it is distributed by the USGS in a different format (USGS 1987).

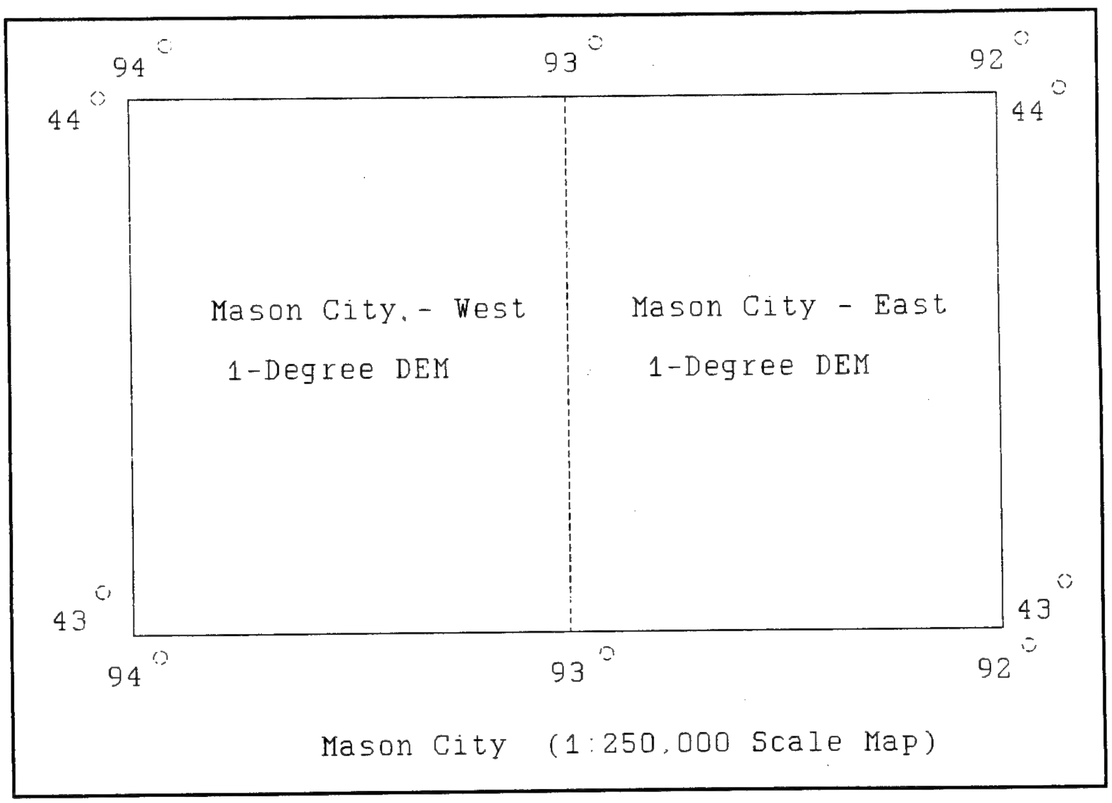



Figure 16. Sample Coverage for 1:125,000 Scale Map and Corresponding 1 Degree DEMs

Data Characteristics

The 1 Degree DEM data consist of a regular array of elevations horizontally referenced to the latitude/longitude geographic coordinate system. Elevation data are provided in meters relative to mean sea level and are rounded to the nearest meter, and the spacing of elevations along each profile in the grid is 3 arc-seconds. Given the spacing of 3 arc-seconds between profiles over a 1 x 1 degree area, the DEM therefore consists of a 1,201 x 1,201 array, or a total of 1,442,201 elevation data points. Since the spacing between points is based on latitude and longitude rather than a rectangular coordinate system, the spacing between points varies by latitude. The East-West spacing between points becomes closer at higher latitudes, while the North-South spacing remains constant. The North-South spacing between points is approximately 92 meters, while the East-West spacing varies from approximately 75 meters in the southern United States to near 65 meters in the Northern areas of the country, not including Alaska. Figure 17 shows an example of the structure for the 1 Degree DEM (USGS 1987).

Data Production

Elevation data for the 1 Degree DEMs are derived from various cartographic and photographic sources. Elevation data from cartographic sources are collected from any map series from 1:24,000 scale through 1:250,000 scale.

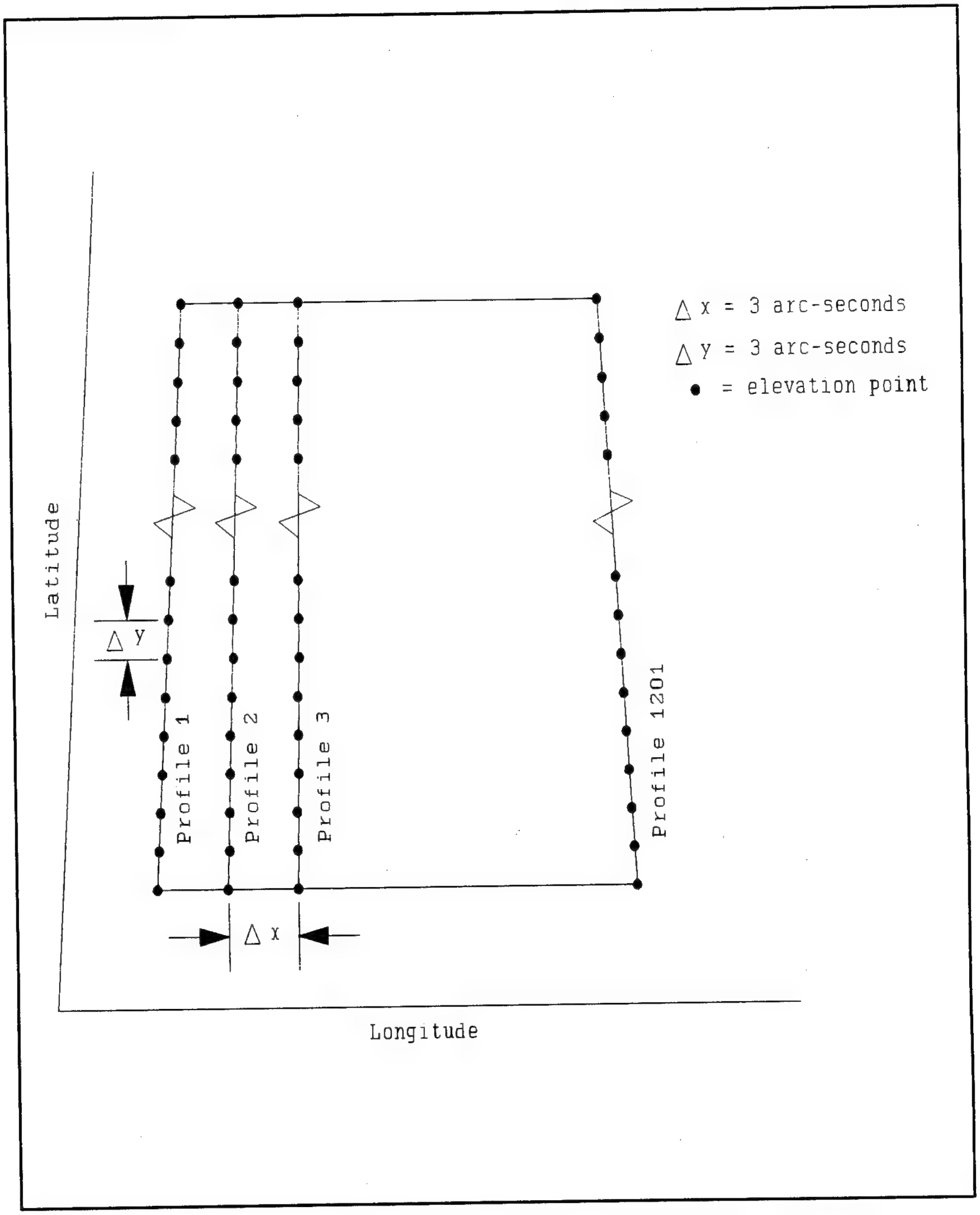



Figure 17. Sample Structure for 1 Degree DEM

Topographic features such as contours, drain lines, ridge lines, and spot elevations are digitized and then processed into the required interval spacing. Data from photographic sources are collected using manual and automated correlation techniques. Elevations are collected along each profile for the DEM, and the raw elevation data are weighted with additional information such as spot heights for final determination of the grid (USGS 1987).

Accuracy

The accuracy of a DEM is dependent on the source of the data and upon the spatial resolution of the data samples. For 1 Degree DEMs, the primary source of data is the 1:250,000 scale maps. The horizontal accuracy of DEM data is primarily dependent upon the horizontal spacing of the grid points. Most of the terrain features are approximated due to the fact that they are reduced to grid nodes spaced at regular intervals. As discussed earlier, this inherently leads to a loss of detail where terrain features do not correspond exactly to the location of grid points. The effect is a smoothing of the land surface during the gridding process because terrain features that are spaced more closely than the DEM grid spacing cannot be represented. The absolute horizontal accuracy (feature to datum) for the 1 Degree DEM is 130 meters circular error at 90 percent probability (USGS 1987).

Vertical accuracy is also dependent upon the spatial resolution, quality of the source data, collection and processing systems, and digitizing systems. The DMA production objective for the 1 Degree data is to satisfy an absolute vertical accuracy (feature to mean sea level) of $\pm$30 meters, linear error at 90 percent probability. Although this accuracy appears to be poor, it is important to note that this is the absolute accuracy for an individual point. For representing terrain to model a watershed, the absolute elevation of each point is not as critical as the relative accuracy between points. For the DEM products, the relative vertical, as well as horizontal, accuracy will in most cases conform to the actual topographic features with greater reliability than would be indicated by the above mentioned absolute accuracy (USGS 1987).

Data Selected

The data selected for use in this research came from the 1 Degree DEM product, available for the entire contiguous United States, Hawaii, and some portions of Alaska. The choice of these data, as opposed to the 7.5 Minute DEM data, was mainly a factor of availability. The 7.5 Minute DEM data are not yet available for the entire United States, while the 1 Degree DEM data are available free via the internet, and are thus readily available to virtually any engineer for modeling any area in the United

States. Although the 7.5 Minute DEM data were not used in this research, the same procedures set forth here for filtering 1 Degree DEM gridded data could also be directly applied to those 7.5 Minute DEM data.

# CHAPTER V

## SELECTION OF DEM POINTS TO FORM A TIN

When presented with the need to generate a TIN from gridded DEM data, the first inclination may be to use the entire DEM grid. After all, when provided with a complete grid of elevation data, why not use all of the grid data? The fact is that many of the elevation points in a DEM grid are totally redundant and serve no purpose in a TIN. For example, DEM data in a flat region may contain dozens or hundreds of adjacent points that are all the same elevation (recall that DEM elevation data are provided to the nearest meter). That same area can be represented in a TIN with handful of points used to form a few large flat triangles. In fact, any area of constant slope in a DEM, no matter how large, can be represented by just a few triangles in a TIN. The challenge then becomes how to select the points in a DEM which are the best points to form a TIN, and to select as few points as possible while still maintaining an accurate representation of the terrain for watershed declination and computation of basin geometric parameters, all of which is the focus of this research. There are several methods for selecting points from a grid to form a TIN, and a discussion of those methods follows.

## Data Point Selection Methods

Several methods have been developed for selecting appropriate data points in an elevation grid for generating a TIN. The "very important point" or VIP method was presented by Chen and Guevera (1987). DeFloriani et al. (1985) detailed the hierarchy method, which is both a data structure for TIN generation and a method for point selection. The drop heuristic method was set forth by Lee (1989) and treats the point selection process as an optimization problem. Finally, there is the method selected for use in this research which is a curvature based filter technique developed by Southard (1990). Each of these methods is discussed below, with a more detailed discussion of the curvature method presented.

### The VIP Method

Chen and Guevera (1987) presented a technique to select "very important points", or VIPs, in a gridded DEM by measuring how well a given grid point is estimated by its eight neighbors. A 3x3 filter defines the neighborhood with the central point in that 3x3 region being the point of interest. The basic assumption of the VIP method is that the greater the difference in elevation between a point and its neighbors, the more important or significant that point will be. This method is carried out by generating four transects through the 3x3 region, one horizontally, one vertically, and two diagonally as shown

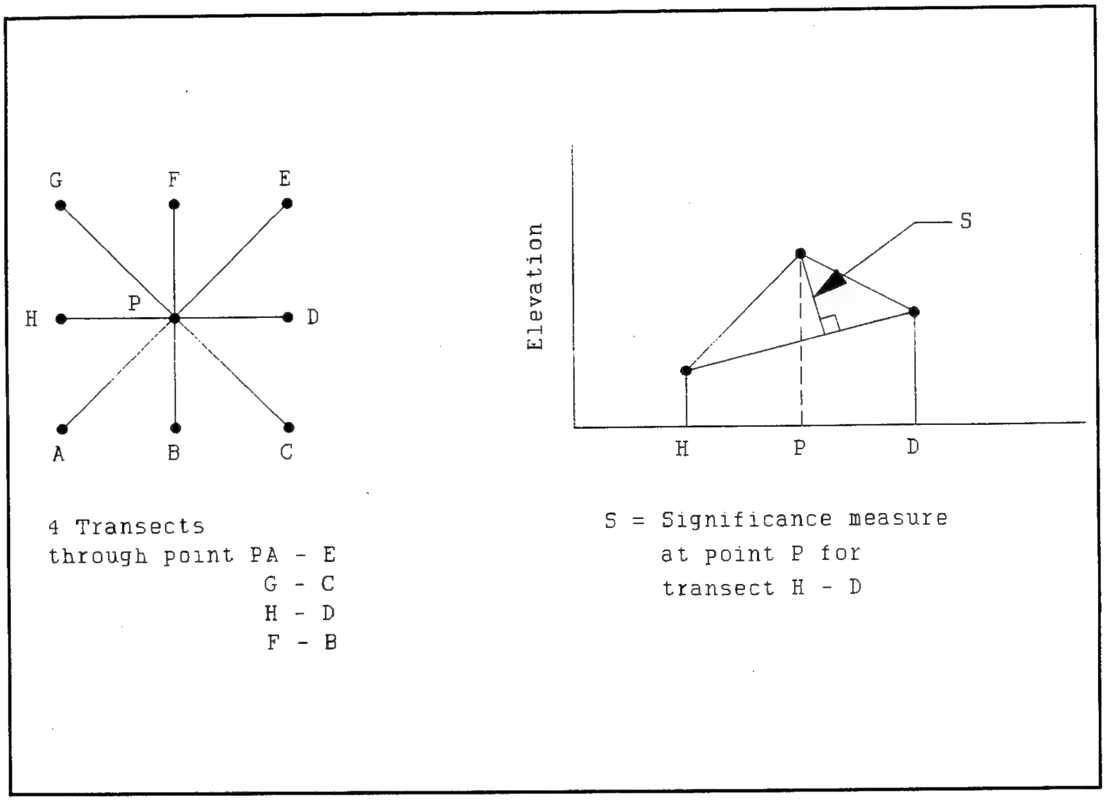



Figure 18. Transects and Significance Value for VIP Method

in Figure 18. For each of those transects the elevation difference between the central point and the transect line is computed, and the absolute value of those four measurements are then summed to determine the "significance value" of the central point. Figure 18 also contains an example of this measurement for one of the four transects. Those points with greater differences in elevation relative to the eight surrounding neighbors will have greater significance values. To determine which points are significant and which are not, the user specifies a particular significance value to serve as a threshold or

specifies a specific number of points to be selected from the grid. As a result of the emphasis on using straight lines, the VIP technique has problems on curved surfaces or in gently sloping topography (Lee 1991b).

### The Hierarchy Method

The hierarchy method, or hierarchy triangulation, is both a data structure and a method by which points can be selected from gridded data to form a TIN. DeFloriani et al. (1985) developed this concept by which the surface is approximated in successively finer levels of detail by triangular patches (Lee 1991a). This algorithm starts by connecting the four corners of the gridded DEM to form two large triangles which encompass the entire DEM grid. Then all elevation points inside each of those triangles are evaluated to compute the difference between their elevation and the elevation at that point as interpolated on the triangle of the TIN. The DEM grid point which has the greatest elevation difference from the TIN is then selected as a TIN vertex and is used to subdivide the larger triangle into smaller triangles. This process is repeated for several iterations, adding a new TIN vertex for each triangle during each iteration. The process stops when the specified number of vertices is reached or when no elevation differences exceed a specified tolerance. Shortcomings of this method are that it tends to produce very long, thin triangles which violate the Delauney

criterion and that it never alters a previous decision during the successive subdivision of triangles. Once an edge has been created, it will remain in the final TIN (Lee 1991b).

### The Heuristic Method

This method, presented by Lee (1989), treats the point selection process as an optimization problem. The method starts with an entire grid DEM and gradually discards the least significant points, one at a time. During each iteration, every remaining point is assessed and the point which causes the least difference in elevation between the grid and the TIN is discarded. The process is stopped either when a particular number of points has been reached or when a specified threshold for elevation differences has been met. This method does utilize the Delauney triangulation such that long, thin triangles are avoided. The heuristic method also assesses the importance of each point in a global context in that it uses the Delauney neighbors instead of the grid neighbors to evaluate the importance of a given point. One problem with this method is that it requires significant computational resources. It evaluates every point remaining in the grid at each iteration, and when the gridded DEM is large the computational time and storage requirements are quite large (Lee 1991b).

## The Curvature Based Filtering Method

The particular method of selecting surface specific points utilized in this research is a curvature based technique proposed by Southard (1990). As with the methods outlined previously, the goal of this method is to select surface specific points from the elevation grid which represent topographically significant points in the terrain from which an acceptable TIN can be constructed. For the purpose of constructing a TIN, the "topographically significant points" that are desired include such features as peaks, ridges, valleys and ravines, passes, pits or depressions, and any other significant break or change in slope. These features comprise the basic set of necessary points from which any terrain surface can be characterized. If these features can be accurately depicted, then the areas between those points can be reasonably modeled as planar surfaces, i.e., as the surface of a triangle in the TIN.

One way to consider each of the features listed above is to recognize that such features create a curvature in the terrain surface, where curvature in this context is defined as a change in slope. Thus, the computation of curvature at each point in the grid is necessary.

### Computing Curvature

The selection of surface specific points in this method is based on an examination of the curvature in the

elevation grid. Second derivatives can be used to detect curvature in the elevation grid describing the terrain surface. In areas with little or no change in slope, curvature values near or equal to zero will result, and these areas can reasonably be represented by the interior planar surfaces of triangles in the TIN. In areas with the most significant changes in slope, the absolute value of the curvature will be greatest. Elevation points with curvature that exceeds a user defined threshold represent points with curvature that is significant, and thus comprise the desired set of surface specific points for generation of a TIN.

To implement this approach digitally, a second order central difference is used to approximate the second derivative at each grid point. In one dimension, the second derivative may be expressed as follows:

$$\left(\frac{\partial^2 f}{\partial x^2}\right)_i = f_{i-1} - 2f_i + f_{i+1} \tag{10}$$

However, in the analysis of terrain features, this must be performed in two dimensions. Similar to the VIP method for evaluating differences in elevation, second derivatives are taken along four transects; namely, one horizontally, one vertically, and one along each diagonal. The resulting filter used in this research for computing a measure of the curvature at each grid point becomes:

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

Note that this will yield a value of zero where there is a constant gradient. For example, take the following cases which each represent a small 3x3 grid with constant gradient. For each case below, the dots represent an elevation point and the numbers near each point represent the elevation of that point. In each of these examples there is a constant gradient and the curvature at the central point in the 3x3 grid will be computed as zero.

Case 1 (flat plane):

$$\begin{matrix} \dot{1} & \dot{1} & \dot{1} \\ \dot{1} & \dot{1} & \dot{1} \\ \dot{1} & \dot{1} & \dot{1} \end{matrix}$$

Case 2 (inclined plane - left to right):

$$\begin{matrix} \dot{1} & \dot{2} & \dot{3} \\ \dot{1} & \dot{2} & \dot{3} \\ \dot{1} & \dot{2} & \dot{3} \end{matrix}$$

Case 3 (inclined plane - upper left to lower right):

| | | |
|---|---|---|
| 1 | 2 | 3 |
| 2 | 3 | 4 |
| 3 | 4 | 5 |

For each of these cases, the curvature, C, is now computed at the central point:

Case 1: $C = (-1*1)+(-1*1)+(-1*1)+(-1*1)+(-1*1)+(-1*1)+(-1*1)+(-1*1)+(8*1)$

$= -1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 + 8$

$= 0$

Case 2: $C = (-1*1)+(-1*2)+(-1*3)+(-1*3)+(-1*3)+(-1*2)+(-1*1)+(-1*1)+(8*2)$

$= -1 - 2 - 3 - 3 - 3 - 2 - 1 - 1 + 16$

$= 0$

Case 3: $C = (-1*1)+(-1*2)+(-1*3)+(-1*4)+(-1*5)+(-1*4)+(-1*3)+(-1*2)+(8*3)$

$= -1 - 2 - 3 - 4 - 5 - 4 - 3 - 2 + 24$

$= 0$

When there is any deviation from a constant gradient, the resultant curvature computed will have an absolute value greater than zero.

Ranking Points

The next step in selecting surface specific points for creating a TIN is to rank them in order of significance

relative to other points in the elevation grid. One way to do so is to select the points with the greatest curvature, regardless of where they fall in the grid. However, this tends to select points only in areas with rapidly changing terrain, such as mountainous regions, and selects few, if any, points in areas with low curvature. The result may be a good model in some areas yet a poor model in others. The methods discussed previously in this chapter use this approach to rank the significance of points, but were based on differences in elevation rather than curvature.

To overcome this problem, Southard (1990) uses a neighborhood ranking scheme such that the curvature of each point is compared to the curvatures of other points in its local region. The local region defined for this research is a 3x3 square shaped neighborhood, with the point of interest being the central point in the 3x3 grid. Larger square neighborhoods, such as 5x5 or 7x7, or circular neighborhoods are possible as well. Once the curvature of each point is computed, each point is then ranked relative to the other points in its neighborhood. The rank is defined as the number of points in the neighborhood whose curvature is less than the central point. Thus, if all other points in the 3x3 neighborhood have a smaller curvature than the central point, then the rank of that central point is 8. Similarly, if no other points in the 3x3 neighborhood have a smaller curvature than the central

point, then the rank of that central point is 0. When applying this method, the user defines the threshold rank, and all points whose rank exceeds that threshold are selected as the surface specific points for creation of the TIN. The effect is that points are selected not only in regions of very high curvature, but locally significant points in areas of lower curvature may also be selected. The selected set of points are thus distributed relatively evenly throughout the area and are not confined only to the more mountainous areas.

Sample Application

A brief example of this method helps to illustrate its application. For this example, a 9x9 grid of elevation points is used. For each of those points, the curvature is computed, the neighborhood rank determined, and the set of surface specific points for various thresholds is selected. Figure 19 shows the sample 9x9 grid, with the x and y axes included for reference. Points in the square 9x9 grid range from (x,y) = (1,1) to (x,y) = (9,9). The number below each point represents the elevation of that point.

The first step is to compute the curvature for each grid point using the procedure outlined above. For example, the point at (x,y) = (3,3) has an elevation of 3, while the eight points in the surrounding neighborhood have elevations of 2, 2, 4, 5, 0, 3, 3, and 3 moving clockwise from the lower left corner of the neighborhood. The

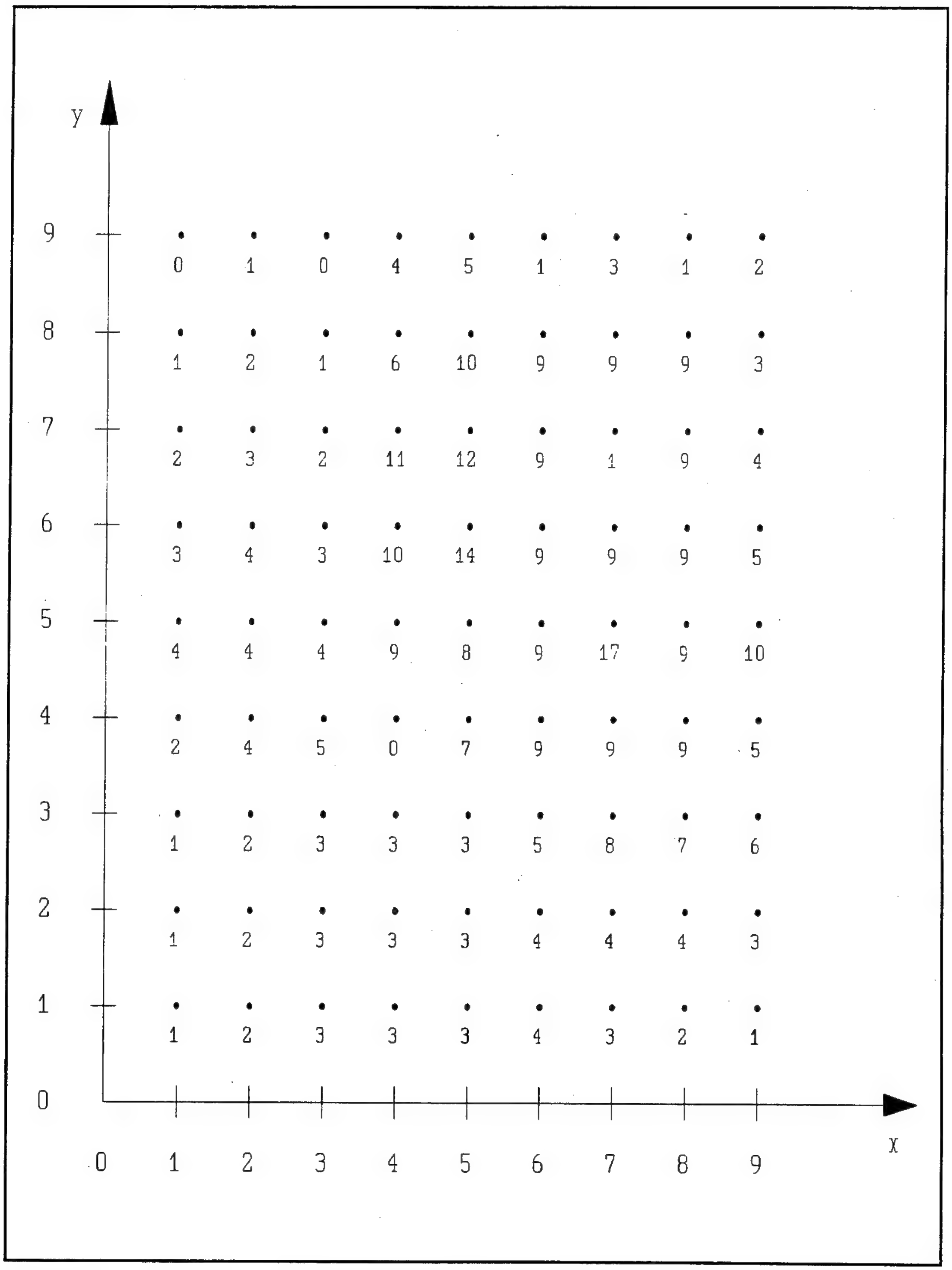



Figure 19. Sample Elevation Grid

curvature of the point located at (x,y) = (3,3) is then computed as follows:

$$
\begin{aligned}
C_{3,3} &= (-1*2)+(-1*2)+(-1*4)+(-1*5)+(-1*0)+(-1*3) \\
&\quad (-1*3)+(-1*3)+(8*3) \\
&= -2 - 2 - 4 - 5 - 0 - 3 - 3 - 3 + 24 \\
&= 2
\end{aligned}
$$

The results of this computation for all data points in this sample grid are provided in Figure 20. Note that upon visual inspection of the original elevation grid in Figure 19, there are several areas that obviously have no curvature. Specifically, the point at location (x,y) = (2,2) is in the center of an inclined plane rising from elevation 1 along the left edge to elevation 3 as it moves from left to right. Additionally, the point located at (x,y) = (4,2) is in the center of a flat plane of elevation 3. The curvature computed at locations (x,y) = (2,2) and (x,y) = (4,2), as shown in Figure 20, is 0 as expected. There are also several locations which should have a high degree of curvature. For example, the point located at (x,y) = (7,7) has an elevation of 1 with all surrounding points at elevation 9. This curvature of that point is -64, the absolute value of which is the highest curvature value in this entire sample grid.

With the curvature computed at each location as shown in Figure 20, the neighborhood rank of each point must then be determined. Recall that this rank is defined as the

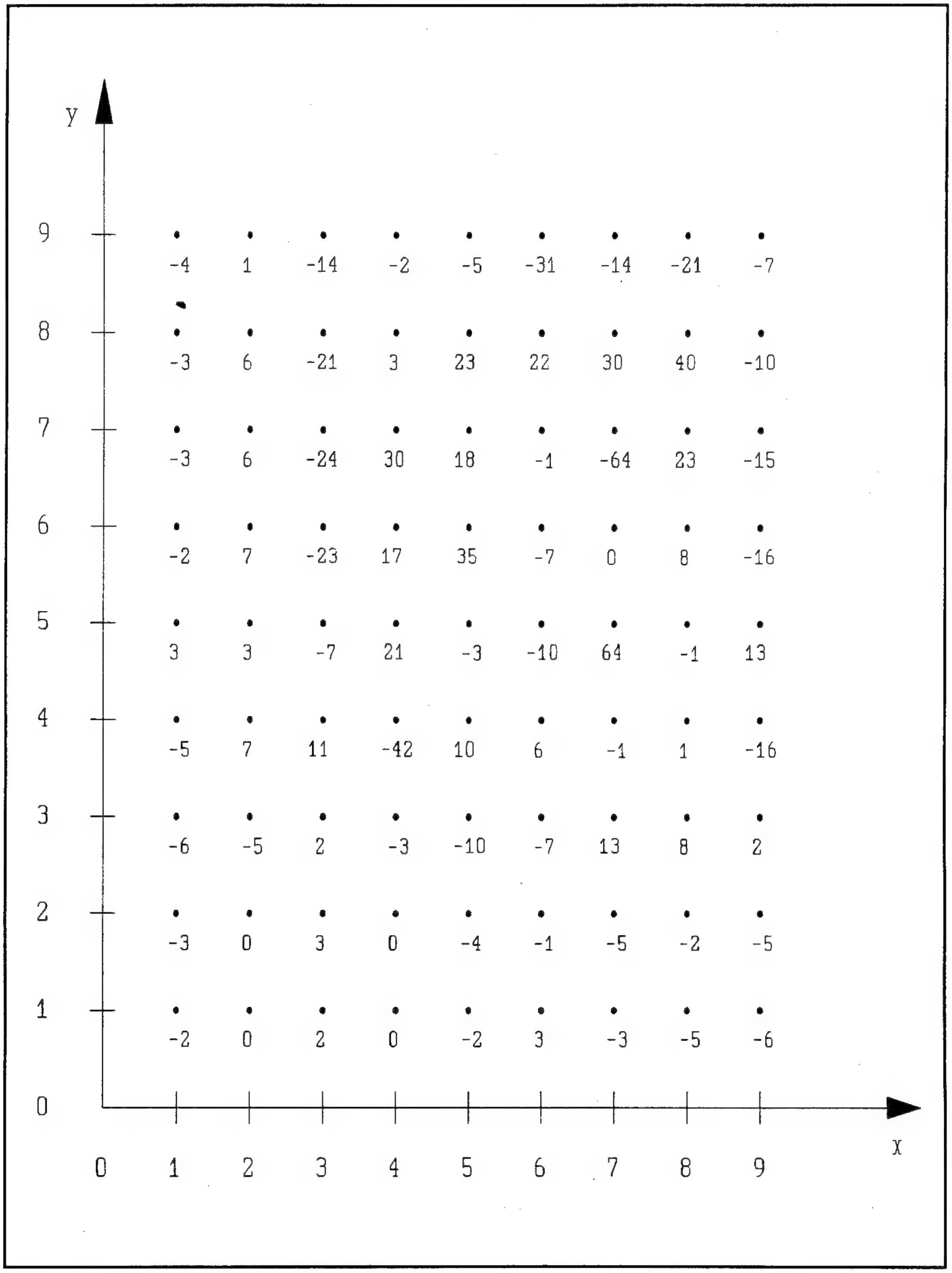



Figure 20. Curvature Values for Sample Elevation Grid

number of points in the 3x3 neighborhood for which the absolute value of the curvature is less than that of the central point. Therefore, a rank of 8 indicates that all eight points in the surrounding neighborhood have a curvature with an absolute value less then the central point, while a rank of 0 means that none of the surrounding points has a curvature less than the central point in the 3x3 neighborhood. For example, recall that the curvature at location (x,y) = (3,3) was computed to be 2. The curvature of the eight points in the surrounding neighborhood, moving clockwise from the lower left corner of that neighborhood, are 0, -5, 7, 11, -42, -3, 0, and 3. Of those eight curvature values, only two of them have an absolute value less than 2, the curvature value of the point at (x,y) = (3,3). Thus, the point at location (x,y) = (3,3) is given a rank of 2. Figure 21 shows the rank of each point in this sample 9x9 grid.

Once the rank of each point is determined, the final step is to specify a threshold for selection of the desired surface specific points. With the technique used in this research, the user specifies a threshold ranging from 0 to 7. Those points with a rank greater than this threshold will be selected as surface specific points from which to create the TIN. For example, if the threshold is 7, then those points with a rank of 8 are retained and all others discarded. Similarly, if the threshold is 3, then all

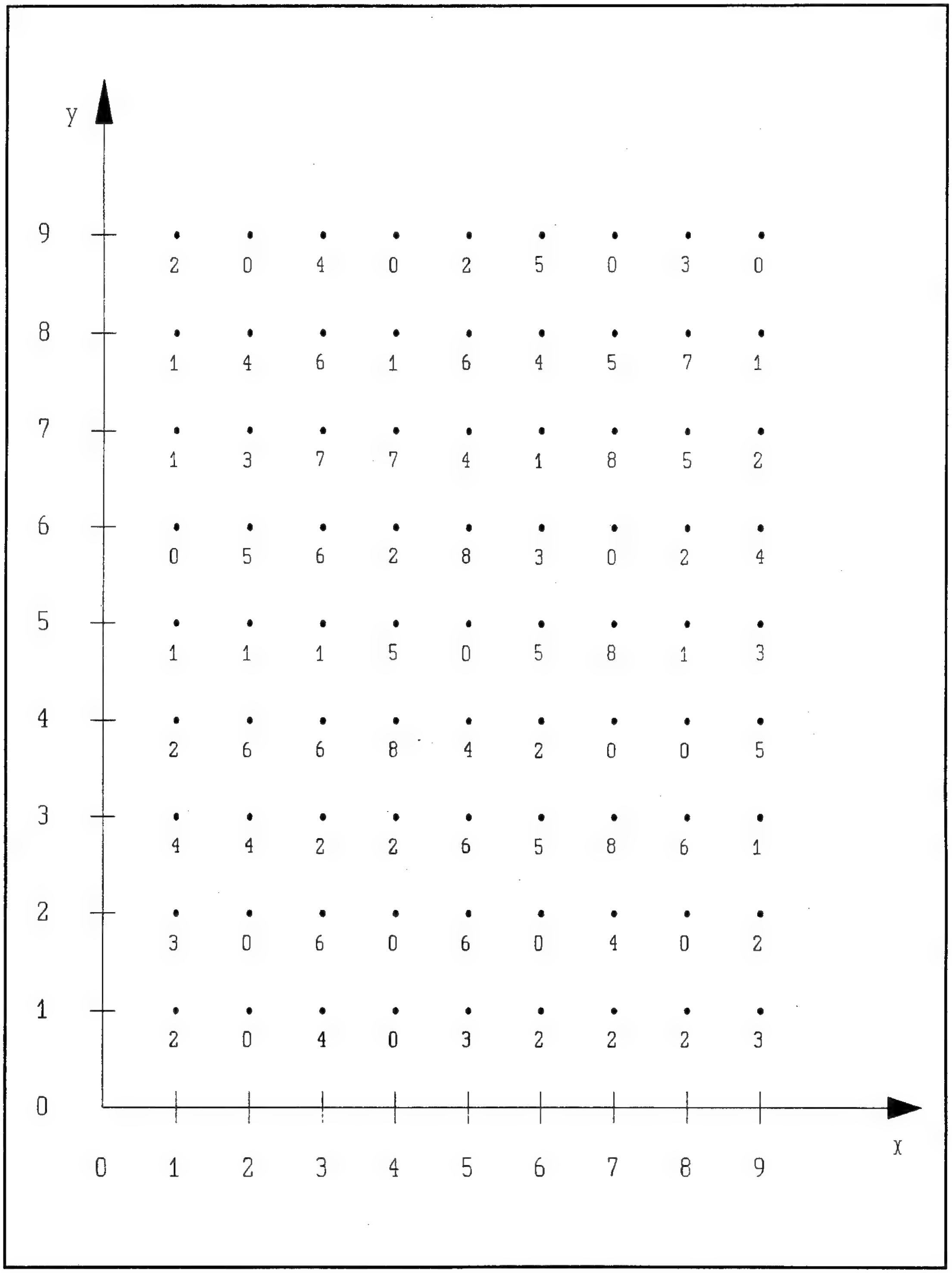



Figure 21. Neighborhood Rank for Sample Elevation Grid

points with a rank of 4 or more are kept while those with a rank of 3 or less are not. For example, Figure 22 includes those points in the sample elevation grid that are retained for creation of the TIN with a threshold value of 7, which includes only those points with a rank of 8. Figure 23 shows those points that would be retained with a threshold value of 3, which includes those points with a rank of 4 or more.

Notice that if the criterion for selection of points did not take the neighborhood rank into account, then the selected points would be different. As shown in Figure 22, for a threshold of 7 there were five points retained. However, those were not the 5 points with the overall highest curvature in this grid. If the criterion for selection was simply to pick the 5 points with the highest curvature, then the points shown in Figure 24 would have been selected. Although this set of points is similar to those with a neighborhood rank of 8 as shown in Figure 22, it is not the exact same set of points. Similarly, as seen in Figure 23, for the threshold value of 3 there were 35 points retained. Once again, if the criterion had been to select the 35 points in the grid with the highest curvature, the result would have been a different set of points as shown in Figure 25. Notice that the neighborhood rank method provides a set of points that are more evenly distributed throughout the grid. This neighborhood ranking

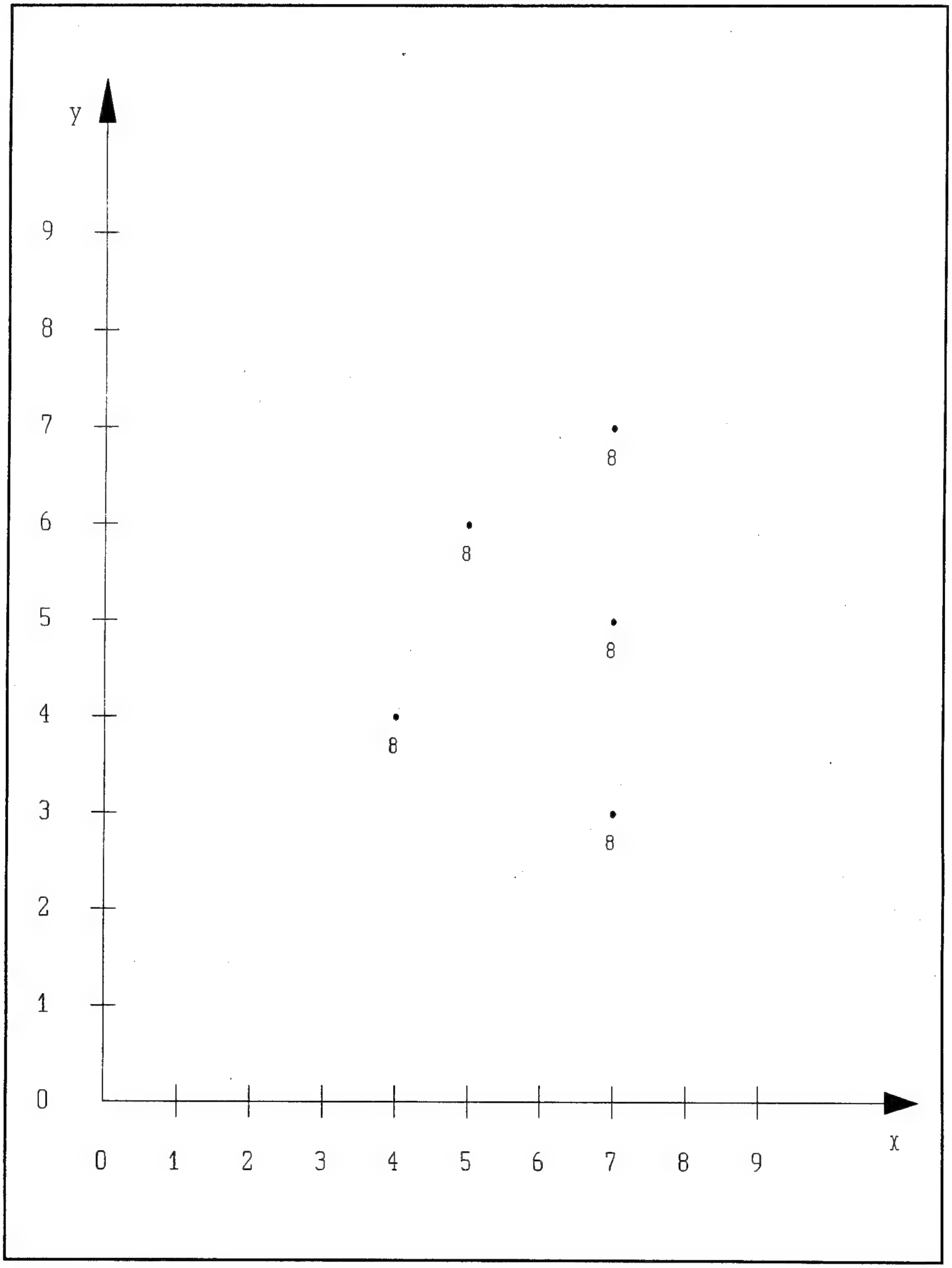



Figure 22. Points in Sample Grid for Threshold Value 7

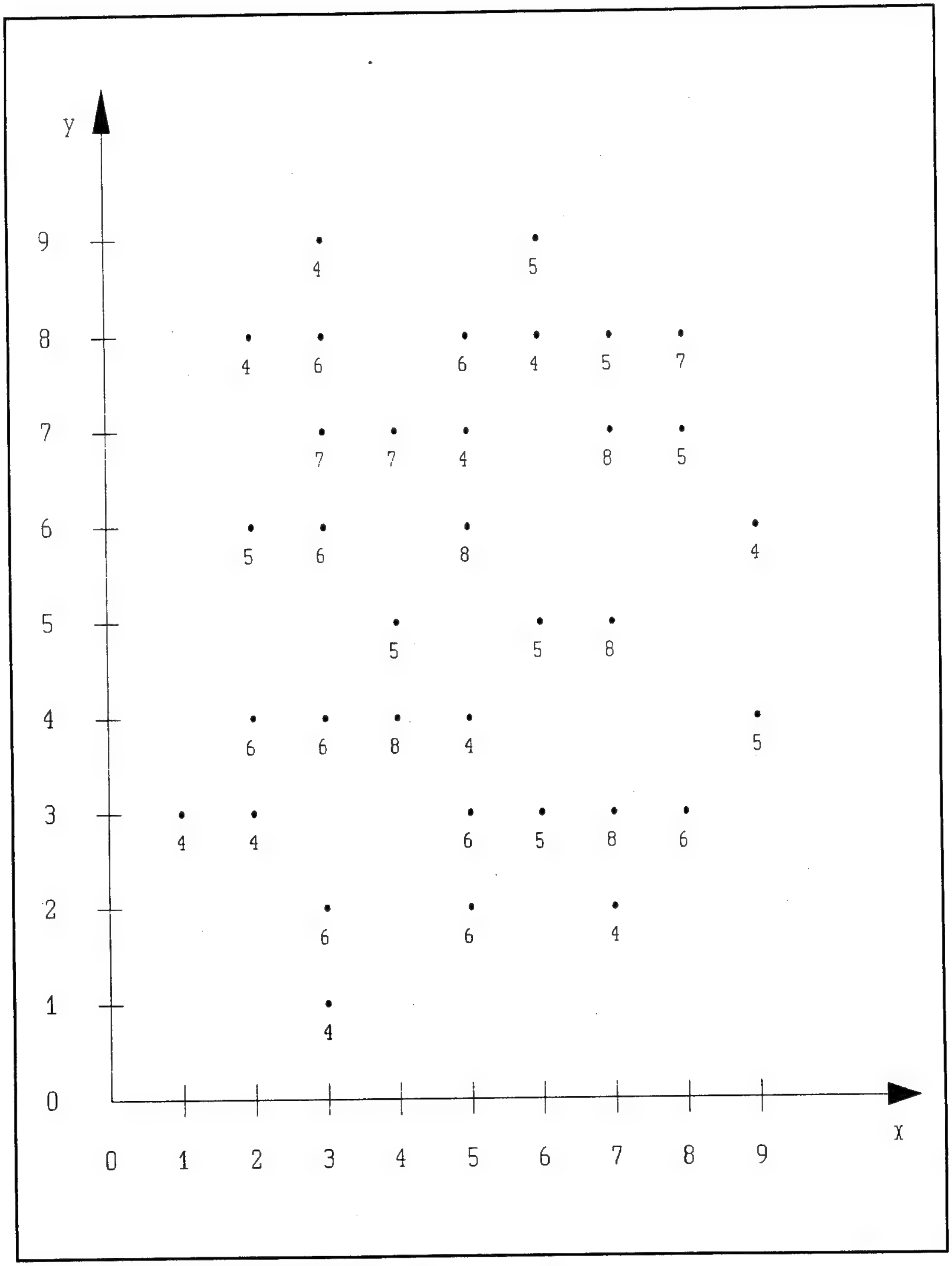



Figure 23. Points in Sample Grid for Threshold Value 3

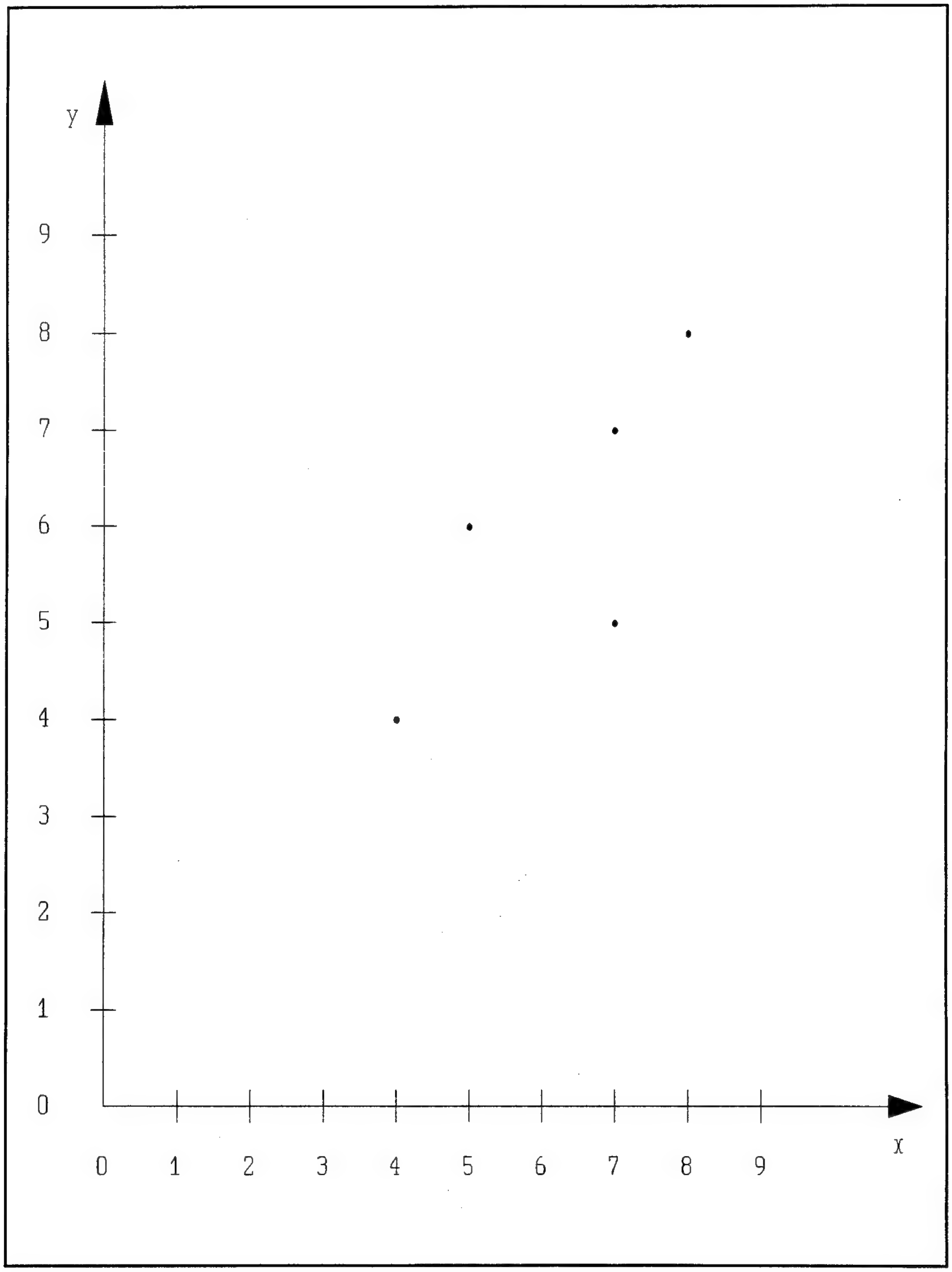



Figure 24. Points with 5 Highest Curvature Values

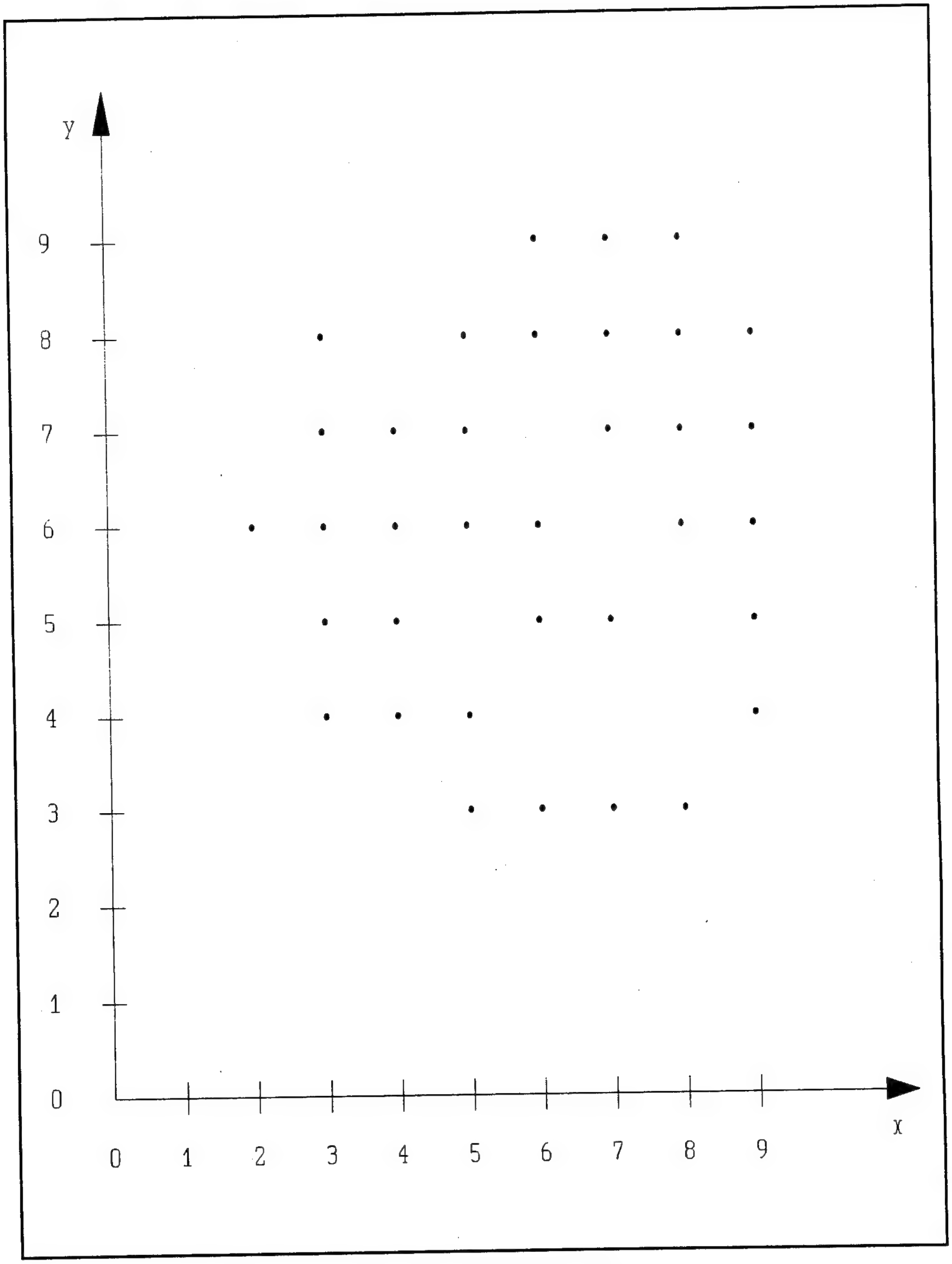



Figure 25. Points with 35 Highest Curvature Values

scheme provides the possibility that points in areas of relatively low curvature globally may be selected if they are locally significant. A good example of this is the grid point located at (x,y) = (7,3) which has a curvature of only 13. There are 23 of the 81 points in this sample grid with a curvature greater than 13, but this point has a neighborhood rank of 8. Although its curvature is not high relative to the entire grid, it is very significant locally and thus is given the highest rank and retained for the TIN.

## Data Filtering Conclusions

Using the curvature based technique for selecting surface specific points outlined in this chapter provides an efficient and logical method of selecting a set of data points from gridded DEM data to create a TIN. This method offers several advantages. It is particularly well suited for selecting points to generate TINs such that the vertices are located where significant changes in slope occur. This method also takes the local significance of points into account, and thus provides a relatively uniform coverage for the entire area being modeled. Another advantage is that the user does not have to specify a particular number of points or a specific tolerance for elevation differences as with the other methods outlined previously. The problem with requiring the user to define either a number of points or a tolerance for elevation

differences is that in order to obtain an acceptable TIN those values may differ significantly depending on the terrain. A particular number of points may generate a good TIN in a relatively flat region, but that same number of points may be completely unacceptable in a mountainous region. With the curvature based method, the user simply defines the threshold value for the neighborhood rank, with the points that exceed that threshold being selected for the TIN. Based upon the specified threshold, this method selects as many or as few points as are appropriate for a given terrain. Of course the selection of an appropriate threshold value is important, and guidance in this selection is a major purpose of this thesis. Once the significant points have been selected though the curvature method, the Delauney triangulation may be used to generate a TIN on which the watershed delineation and geometric computations outlined in Chapter III can be performed.

# CHAPTER VI

# STUDY WATERSHEDS AND DATA FILTERING

The goal of this research was to analyze the impacts of the curvature based filtering process on the generation of TINs from gridded DEM data and to provide guidance on acceptable levels of filtering. There are several ways in which those impacts can be assessed. One way is to review the statistical error obtained from various degrees of filtering, whereas, another is to look at the impact of data filtering on watershed delineation and the computation of watershed geometric attributes. Finally, the impacts of using in a hydrologic model the watershed geometric attributes resulting from different levels of filtering can be assessed. The remainder of this chapter discusses the watersheds used in this study and the filtering of DEM data on those watersheds. The following three chapters are then dedicated to each of the ways noted above of assessing the impacts of the data filtering.

## Study Watersheds

To fully analyze the impacts of the curvature based data filtering process on DEM data, different types of topography were considered since different results could be

expected in different types of terrain. Relatively flat areas with low curvature values will undoubtedly respond to the filtering process in a different manner than will a highly mountainous region with more extreme curvatures. Thus, watersheds in three different types of terrain were chosen for this study. Each of the watersheds is comparable in size, but each is representative of a different type of terrain. Following is a brief description of these three watershed areas.

## Rock Creek Watershed

The Rock Creek Watershed is located in steep mountainous terrain in central Arizona. Rock Creek flows into Sycamore Creek near Phoenix, Arizona, and is a tributary to the Verde River. This watershed covers approximately 65 $km^2$ in area, has an average slope near 0.20, and has an elevation change of 1,270 meters ranging from the maximum of 1,900 meters to the low point at 630 meters above mean sea level. The 1 degree DEM elevation data for that region is contained in the Mesa West DEM, and Figure 26 provides a general vicinity map for the Rock Creek Watershed.

## Long Creek Watershed

The second watershed used in this study was a portion of the Long Creek Watershed in northern Mississippi. Long Creek is located in gently rolling terrain near Batesville,

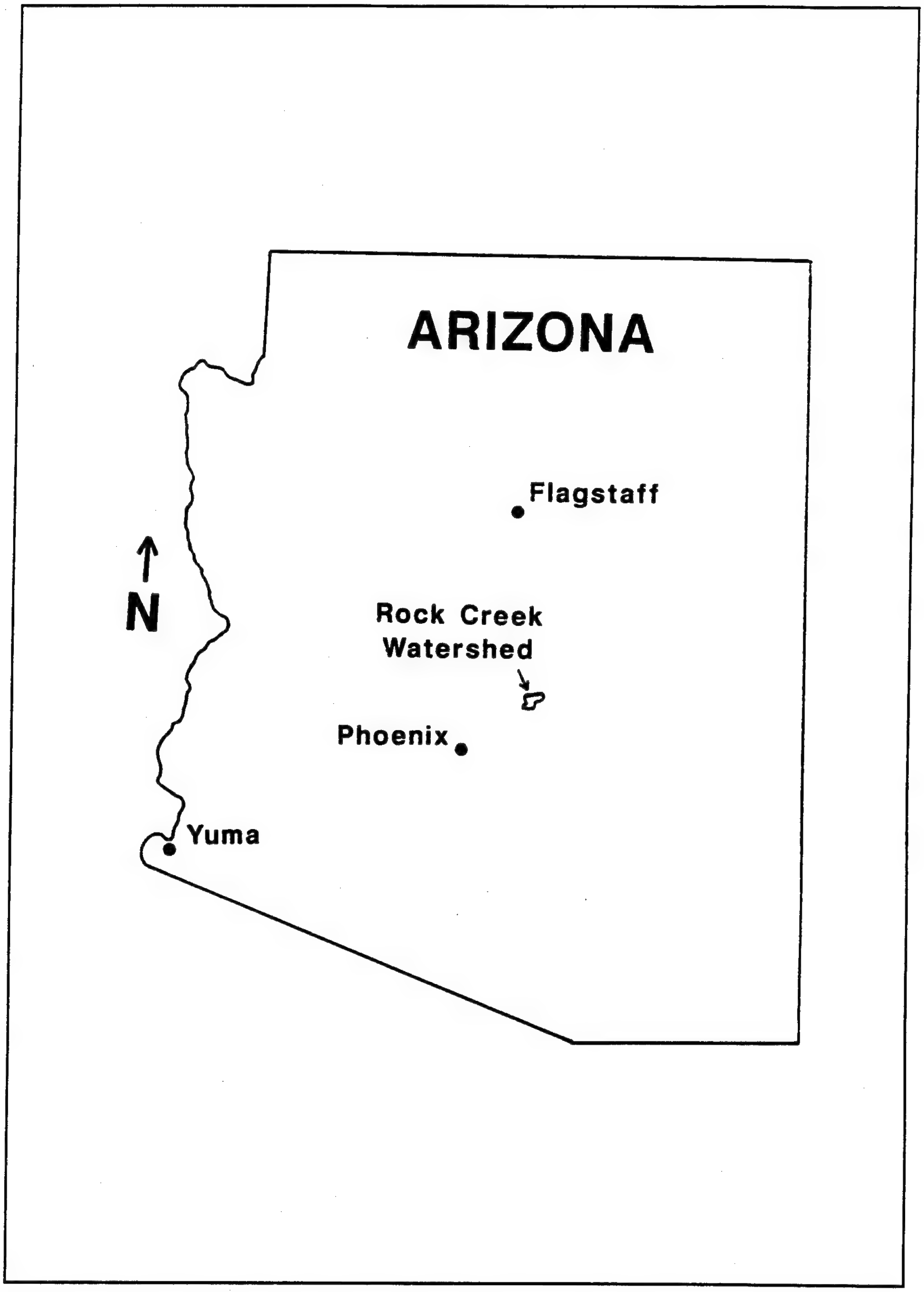



Figure 26. General Vicinity Map for Rock Creek Watershed

Mississippi, and ultimately flows into the Tallahatchie River. The portion of the Long Creek Watershed used in this study has an area of approximately 80 $km^2$, has an average slope of roughly 0.012, and includes a change in elevation of about 75 meters ranging from 150 meters to 75 meters above mean sea level. Elevation data for this area was obtained from the Tupelo West DEM. Figure 27 is a general vicinity map for the Long Creek Watershed.

### Beaverdam Creek Watershed

The third area used in this study is a portion of the Beaverdam Creek Watershed located in a very flat region of northern Iowa near Mason City, Iowa. Beaverdam Creek is located in the upper reaches of the West Fork Cedar River, a tributary to the Cedar River which flows into the Shell Rock River. The Beaverdam Creek Watershed area is approximately 42 $km^2$, has an average slope of about 0.006, and has an elevation difference of about 70 meters ranging from the high elevation of 395 meters down to 325 meters above mean sea level. The Mason City West DEM contains elevation data for this area, and Figure 28 includes a general vicinity map for the Beaverdam Creek Watershed.

## Data Filtering

The first step in performing the analyses on the curvature based data filtering process was to preprocess the DEM data. As stated earlier, each USGS 1 Degree DEM

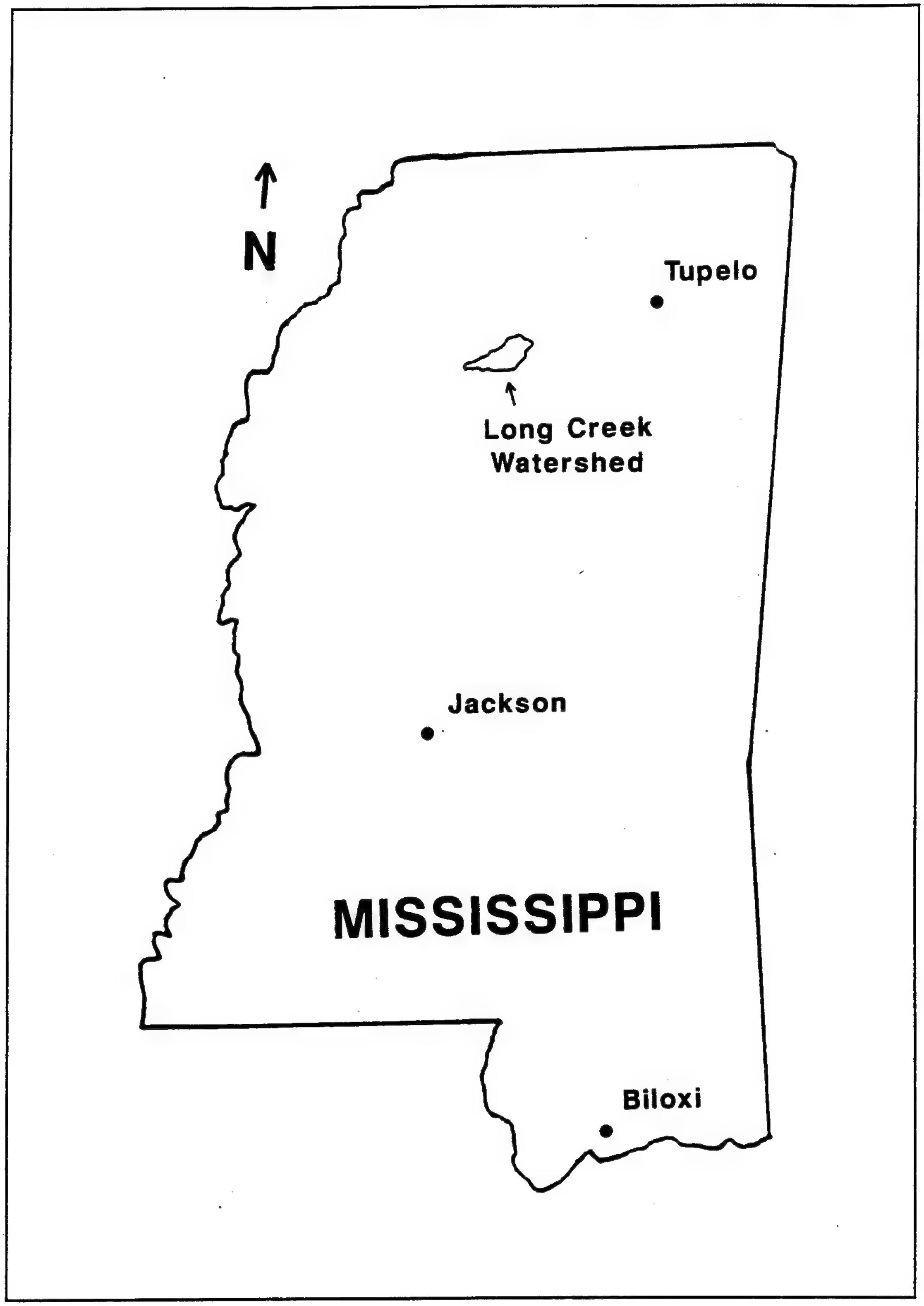



Figure 27. General Vicinity Map for Long Creek Watershed

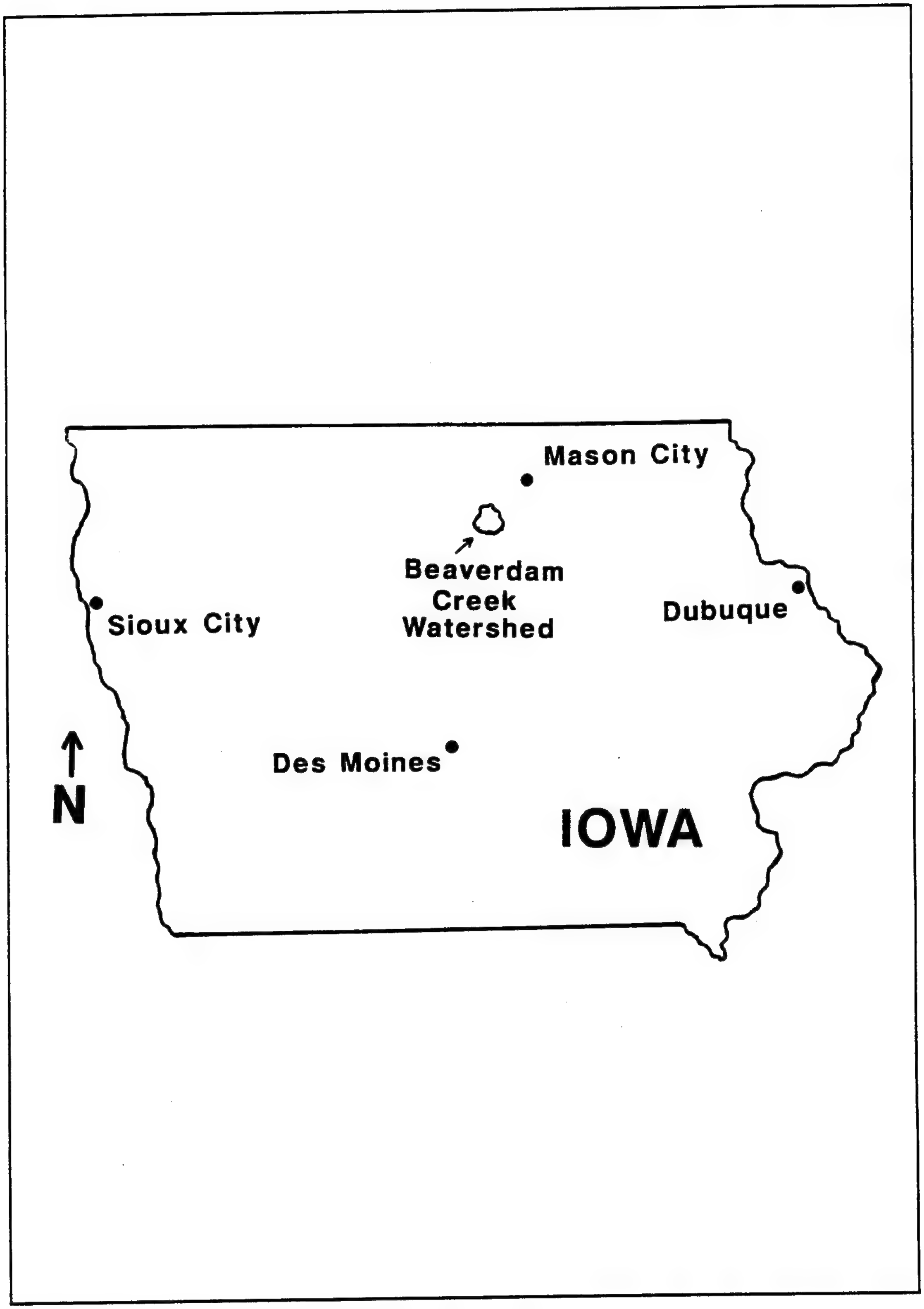



Figure 28. General Vicinity Map for Beaverdam Creek Watershed

product covers a 1 x 1 degree area and contains 1,442,401 elevation points. The total area covered by each 1 Degree DEM is approximately 9,000 $km^2$, depending on the latitude of the DEM. The watershed areas used in this research are all much less than 100 $km^2$ in area and cover only relatively small portions of 1 x 1 degree areas. Therefore, only small portions of each DEM were needed. To establish workable data sets, rectangular areas which encompass each watershed were extracted from each DEM. A rectangular region of the Mesa West DEM which encompasses the Rock Creek Watershed contained 201 points in the east-west direction and 101 points in the north-south direction. Therefore, the DEM grid for Rock Creek consisted of a total of 20,301 elevation points. For Long Creek, the rectangular region of the Tupelo West DEM encompassing the watershed contained 241 points east-west and 141 points north-south for a total of 33,981 points in the Long Creek DEM grid. The DEM grid for Beaverdam Creek included 131 points east-west and 213 points north-south from the Mason City West DEM, for a total of 27,903 points. These three DEM grids were then used in the filtering process as detailed below.

With the DEM grids established for each watershed area, the next step was to filter the DEM data using the curvature based method outlined in Chapter V. The DEM data for each of the three watershed areas was filtered at all

eight of the possible filter threshold levels, which resulted in a total of 24 data sets. Due to the nature of the filtering process, a different percentage of points at each filter threshold was retained for each of the three areas. That is, a filter threshold of 7 produces a different percentage of data filtering in a flat region than in a mountainous region. The results of the filtering process are summarized in Table 1.

The Rock Creek area has much less flat terrain than the other two, which is to be expected since the Rock Creek Watershed is the most mountainous of the three regions. This can be seen by looking at the 0 threshold level data in Table 1. Recall that a filter threshold of 0 will filter out only those points which have a rank of 0, which are points where all eight of the surrounding neighbors have a greater curvature value. In flat terrain, there will be a higher percentage of points with a rank of 0 than in mountainous areas. This is reflected in Table 1 in that only 18.3% of the points in the Rock Creek area had a rank of 0 while 38.2% of the points in the Beaverdam Creek region had 0 rank. It is interesting to note that while the variation in the amount of data filtered is fairly pronounced at the low filter threshold, at the highest threshold level the percentage of data points removed by the filtering is almost constant at about 95%. This indicates that each of the three study areas has

Table 1. Summary of DEM Data Filtering

| Summary of DEM Data Filtering | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| ROCK CREEK (Original # of Points: 20301) | | | | | | | | |
| | Filtering Threshold | | | | | | | |
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| # of Pts. Rmvd | 3720 | 5606 | 7588 | 9754 | 12182 | 14599 | 16859 | 18824 |
| % of Pts. Rmvd | 18.3 | 27.6 | 37.4 | 48.0 | 60.0 | 71.9 | 83.0 | 93.0 |
| LONG CREEK (Original # of Points: 33891) | | | | | | | | |
| | Filter Threshold | | | | | | | |
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| # of Pts. Rmvd | 11860 | 14534 | 17121 | 20105 | 23712 | 26917 | 29698 | 32205 |
| % of Pts. Rmvd | 35.0 | 42.9 | 50.5 | 59.3 | 70.0 | 79.4 | 87.6 | 95.0 |
| BEAVERDAM CREEK (Original # of Points: 27903) | | | | | | | | |
| | Filter Threshold | | | | | | | |
| | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| # of Pts. Rmvd | 10761 | 12447 | 14144 | 17143 | 21083 | 23188 | 25058 | 26849 |
| % of Pts. Rmvd | 38.2 | 44.6 | 50.7 | 61.4 | 75.6 | 83.1 | 89.8 | 96.2 |

approximately the same percentage of points with a rank of 8, which again are points where the curvature of all eight surrounding neighbors is less than that of the central point.

While the contents of Table 1 provide a summary of the number of points and percentage of points removed from each data set for each filtering threshold, that information does not provide any indication on the spatial distribution of the points retained. An actual plot of the points retained in each data set is very useful in understanding and visualizing the results of this filtering process. Appendix A contains plots of the points retained for each data set. Each of the plots in Appendix A depict the actual spatial distribution of the data points retained for each filtering threshold. There are 9 plots provided for each geographic area represented in this study. The first plot for each area is the DEM grid before any filtering was performed, and the following 8 plots for each area depict the results of filtering thresholds 0 through 7.

## TIN Generation

The focus of this research is to determine the effects of this filtering process on the geometric data derived from TINs that are subsequently used in hydrologic modeling. Thus, the next step in the processing of the elevation data was to generate a TIN from each of the sets of filtered data presented in Appendix A for the three

geographic regions. Following the Delauney criterion, a TIN for each of these 27 sets of elevation data were created. Appendix B contains plots of a portion of each of these TINs. The density of lines and points on most of these TINs is such that a plot of the entire TIN on a sheet of paper is nothing more than a blackened area with a few small white regions where points were filtered out of the grid. Therefore, only a small region from each TIN, approximately one-fourth of each entire TIN area, is presented in the plots in Appendix B. Although the entire TIN is not shown, these plots do provide a clear view of the TIN structure and give a good indication of the relative density of points that resulted for each filtering threshold.

The TINs generated during this process were then analyzed to assess the quality of the TIN, as compared to the original DEM grid. The results of watershed delineation and computation of basin geometric parameters using the TINs generated from data at various filtering thresholds were also analyzed. The following chapters detail these analyses.

# CHAPTER VII

# STATISTICAL ERROR ANALYSIS

Given the TINs that have now been generated at each filtering threshold for the three types of terrain, some measure of the quality of those TINs is desirable. A terrain surface represented by a TIN generated from the entire original DEM grid, with no data filtered out, would provide a terrain model basically equivalent to that original DEM grid. However, as some data points are filtered out of the original DEM grid and a TIN is created using only some portion of those original DEM points, then the surface represented by that TIN will deviate from the DEM to some degree. The focus of this chapter is to provide some measure of the deviation between the TINs that have been generated from the various levels of filtered DEM data and the original DEM surface.

## Error Computation

There are several ways in which the surfaces represented by a TIN and by the original DEM data can be compared. Each TIN which is generated from some portion of the original DEM data is essentially an approximation of the surface represented by that DEM. The method chosen to

measure the difference between the various TIN and DEM surfaces in this research is to compute the difference in elevation between the TIN surface and the DEM at each of the original DEM grid point locations. For example, the original DEM grid for the Rock Creek area contained 20,301 data points, and for the filter threshold of 6 a TIN was created using only 3,442 of those points. The first step in comparing those two surfaces was to compute the difference between the original DEM elevation value and the elevation value on the TIN at each of the original 20,301 locations. For this example, the result would be a data set consisting of 20,301 error values.

The first step in this process is to interpolate an elevation for any given x,y location on a TIN. This is needed so that for each original DEM grid point location, the elevation on the TIN at that x,y location can be compared to the original DEM elevation value. This process involves several steps for each DEM grid point. The first step for each given x,y location is to determine which triangle in the TIN that x,y location falls within. Once the proper triangle has been identified, then the equation which defines the plane on which that triangle lies must be computed. With that plane equation, the elevation can then be interpolated for the x,y location.

A computer program was written to perform these computations. The program first reads in all of the x,y,z

vertex data and the triangle data from the TIN file (recall the TIN data structure from Chapter III). The program then reads in the first grid point x,y location from the DEM file and loops through the triangle data to determine which triangle that x,y location falls within. With the appropriate triangle identified, the equation of the plane which is defined by the three vertices of the triangle is determined, and the elevation at the given x,y location on that plane is then interpolated. That TIN elevation is then compared to the original DEM elevation to determine the difference. This process is then repeated for each of the remaining x,y grid locations in the DEM file. The output from this program is a file which contains an error value for each DEM grid location, where the error here is defined as the difference in elevation between the DEM and the TIN at a given location. This process was performed on each of the 24 TINs that were generated for this research, one for each of the 8 levels of filtering in each of the three regions.

## Error Distributions

Prior to computing any statistical measures for the error data, it is often helpful to look at the distribution of those data. Figures 29, 30, and 31 provide a graphical depiction of the error distribution, or a histogram plot, for each of the watershed areas. These plots indicate the error values versus the number of data points with that

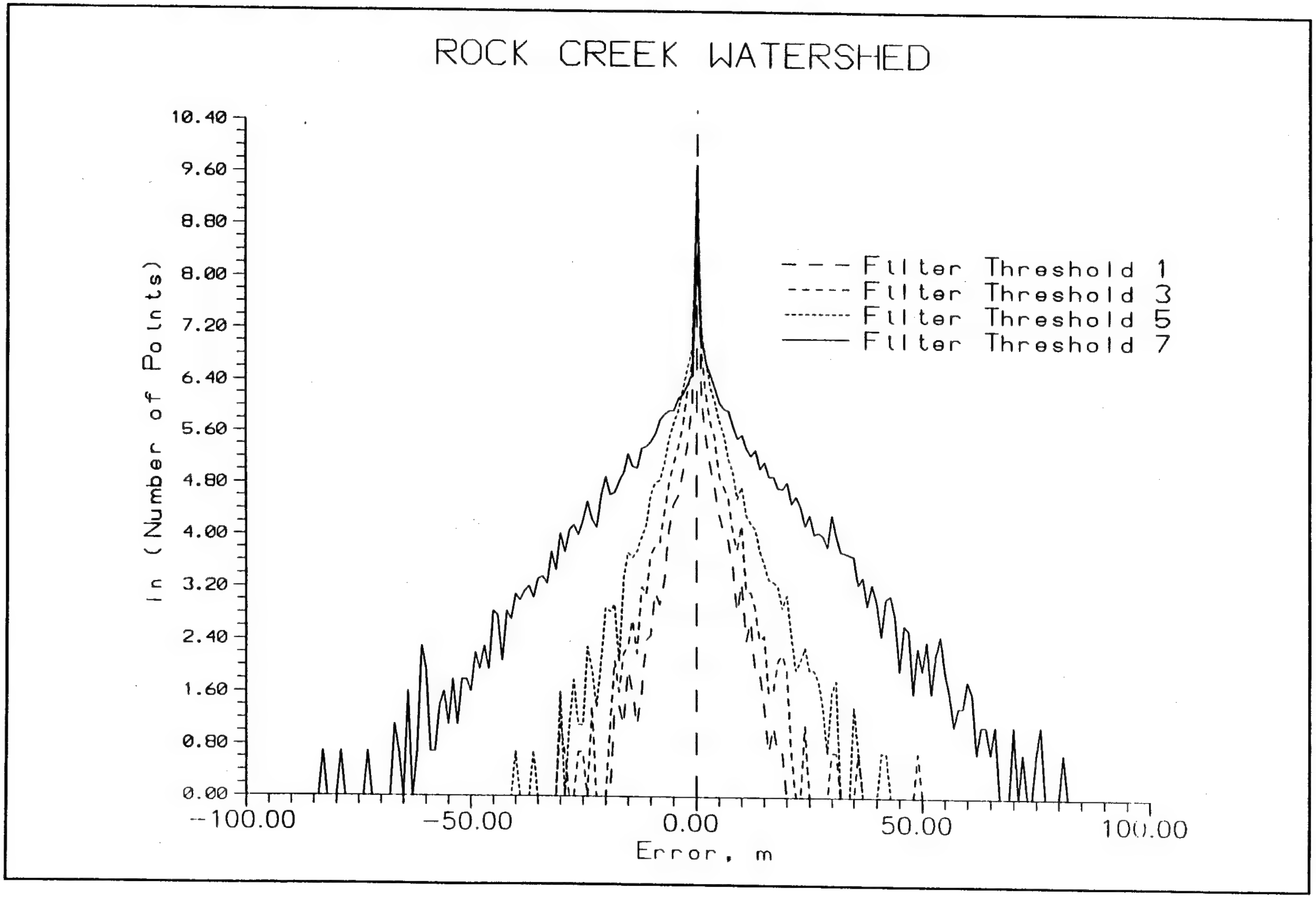



Figure 29. Histogram Plot for Rock Creek Error Data

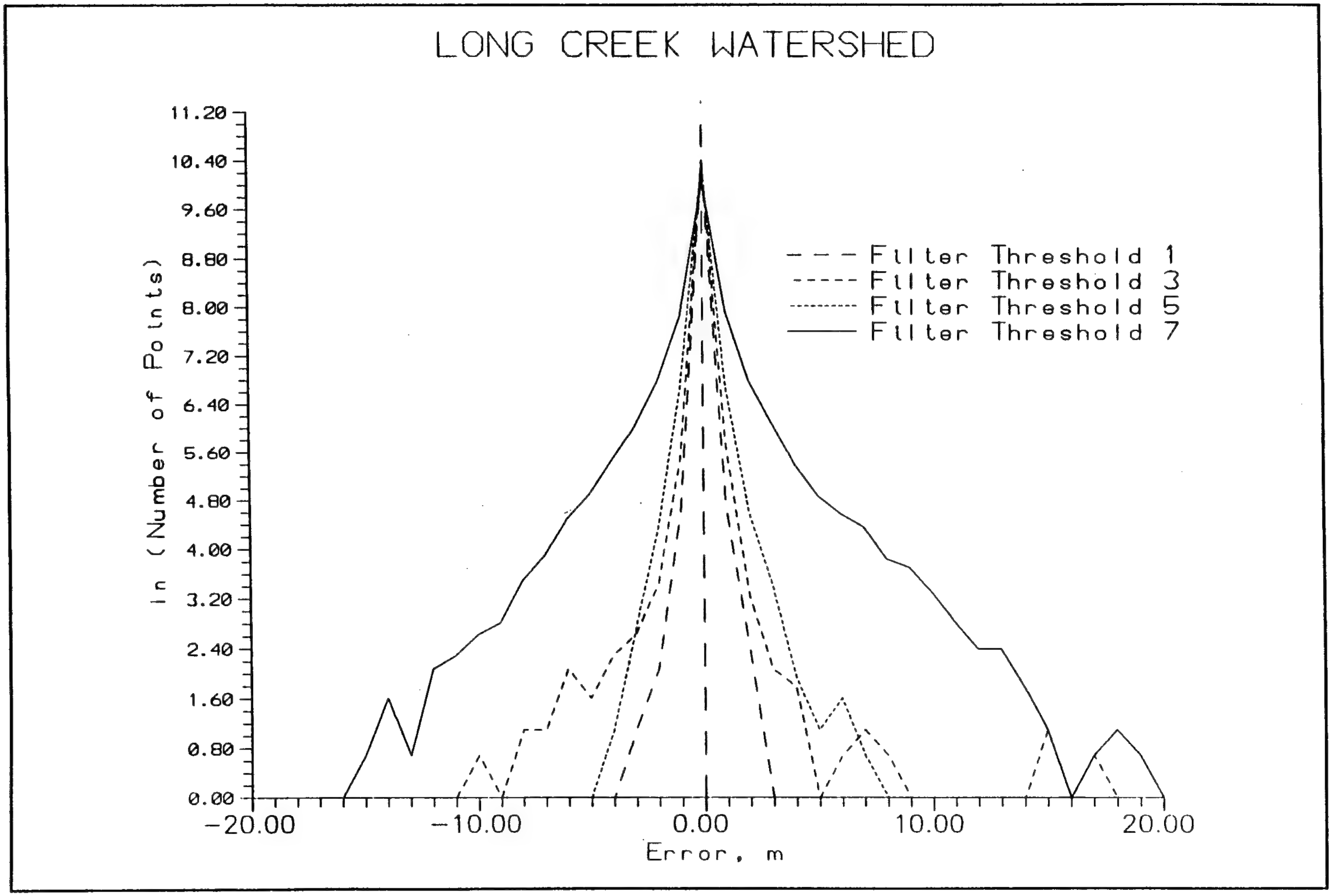



Figure 30. Histogram Plot for Long Creek Error Data

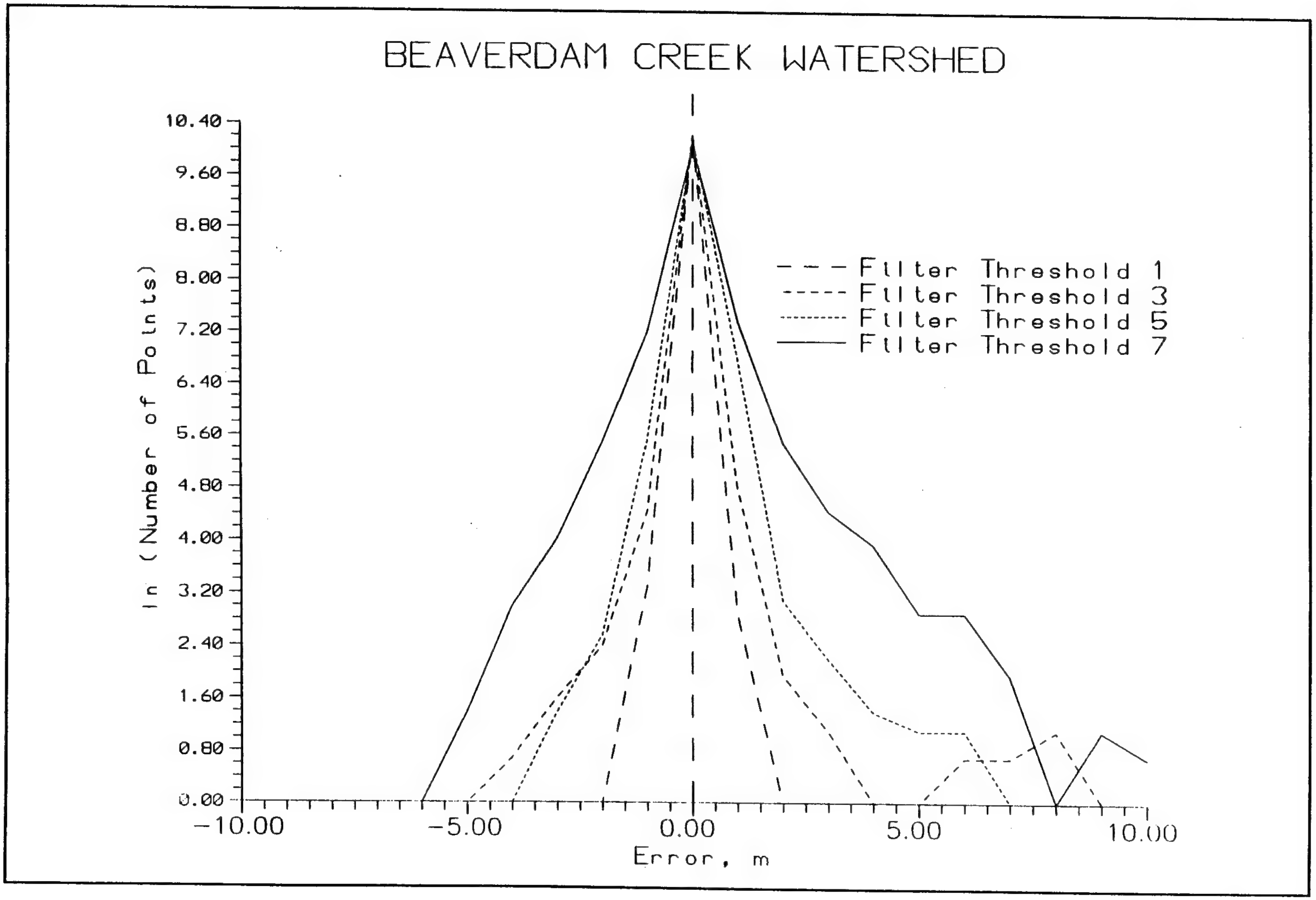



Figure 31. Histogram Plot for Beaverdam Creek Error Data

error value. Since the elevation on the TIN may either be above or below the elevation at that location on the original DEM, then the error at each point may either be greater than or less than zero. The error is distributed around 0, and as more data are filtered from the DEM the number of points with larger error values increases. This results in the "spreading out" of the plots in Figures 29, 30, and 31 as the filter threshold increases. If a particular TIN surface provides a perfect representation of the original DEM data, then the error histogram plots would be a thin spike located at an error value of 0. Obviously, there is error in these TINs, and that error increases as the filter threshold increases and more data are filtered out of the DEM file. The vast majority of the error values in each of the data sets are at or very near zero, particularly for the lower filter threshold levels, and there are only a relatively few points with a large error. To make the plots in Figures 29, 30, and 31 more readable, the vertical axes indicate the natural logarithm of the number of points, and the error data are plotted for only 4 of the 8 filtering thresholds.

Another way in which the error data can be viewed is to plot the percentage of points that exceed certain error values, as shown in Figures 32, 33, and 34. The measure of error that is of interest here is the distance between the TIN surface and the DEM grid elevation at each grid

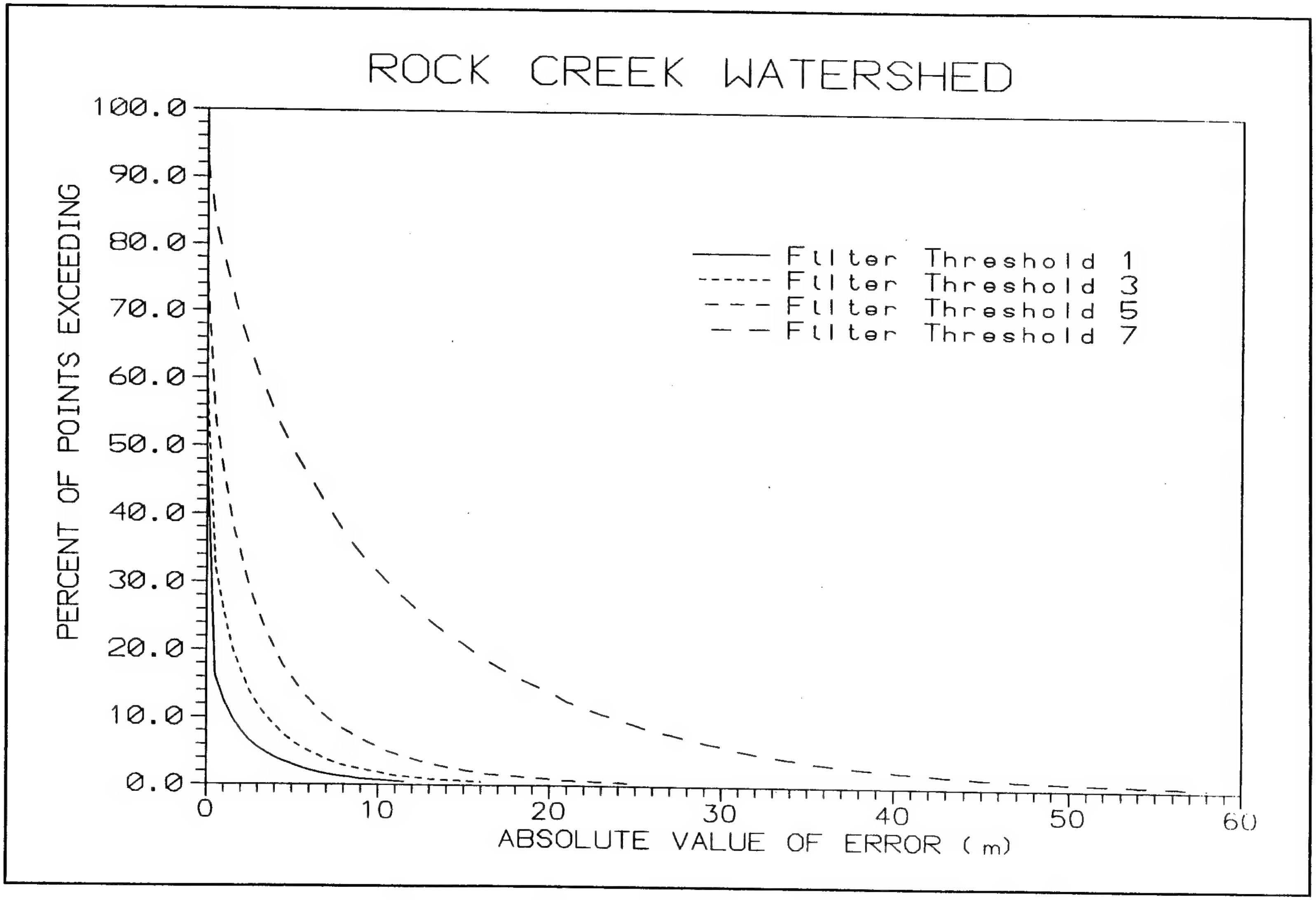



Figure 32. Percent of Points Exceeding Error for Rock Creek

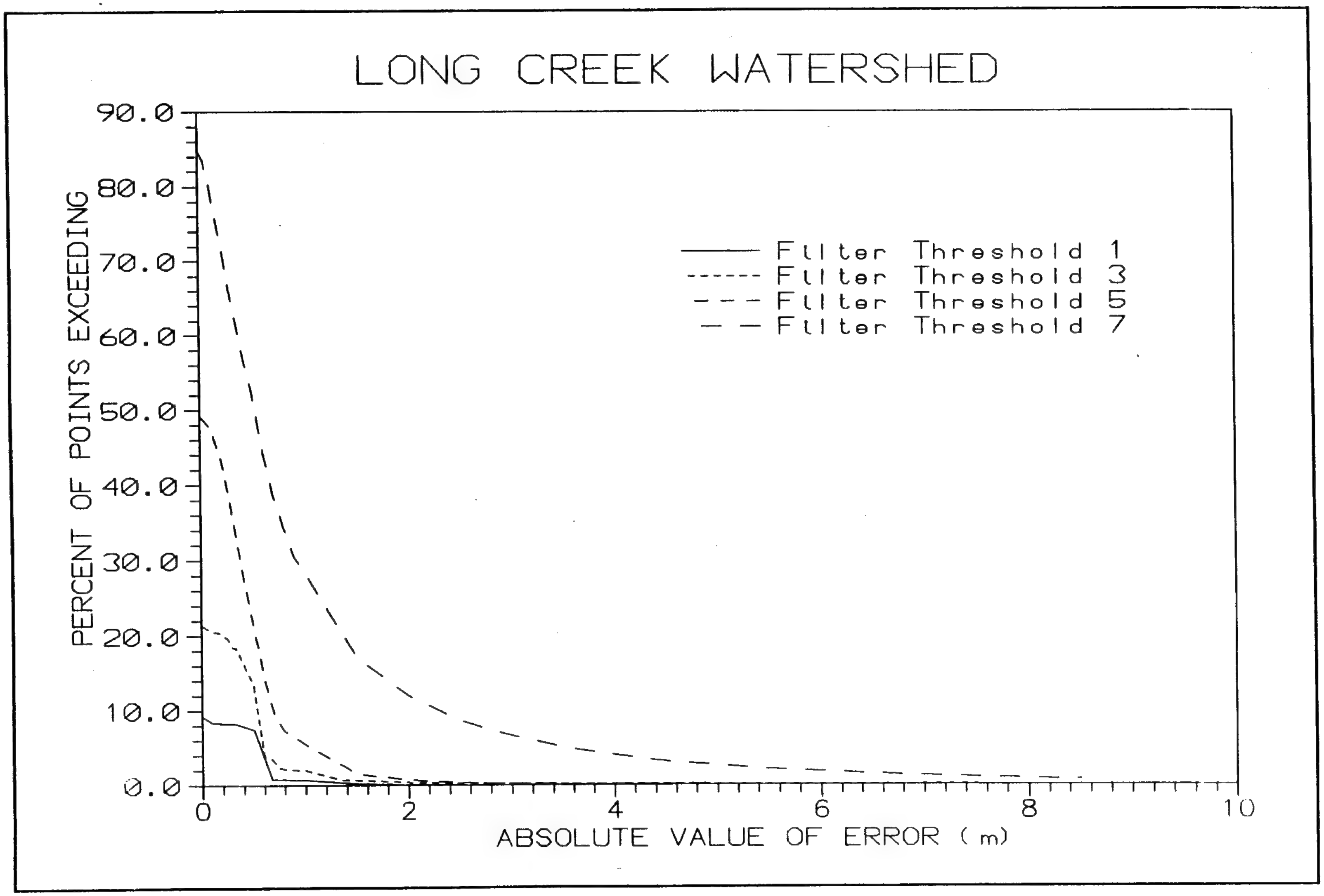



Figure 33. Percent of Points Exceeding Error for Long Creek

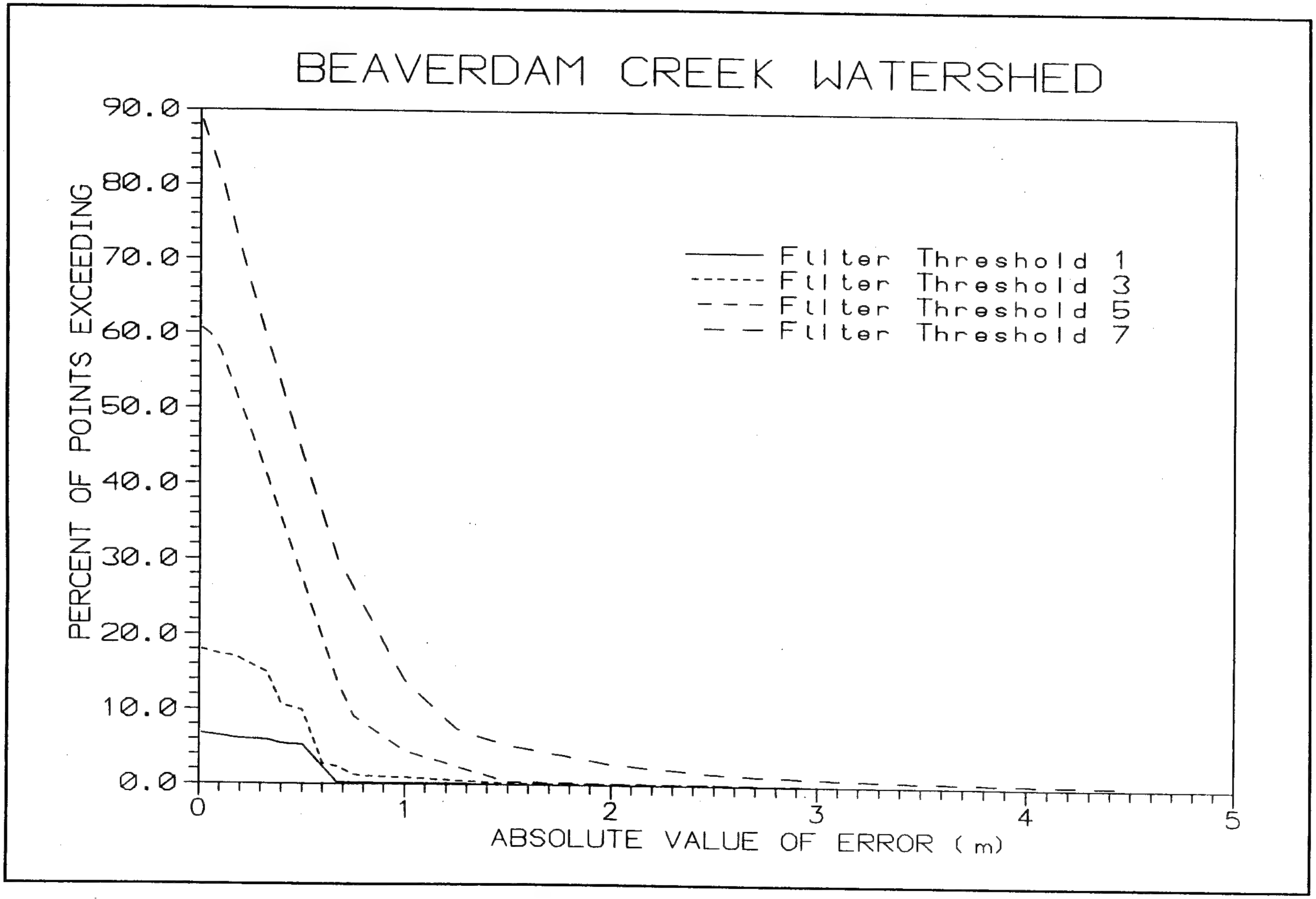



Figure 34. Percent of Points Exceeding Error for Beaverdam Creek

location, and the absolute value of the error provides this measure. Therefore, in Figures 33, 33, and 34 the absolute value of the error is plotted against the percentage of points that exceed that error.

## Statistical Measures of Error

Statistical analysis was performed on the error data to summarize the maximum error, the mean of the absolute value of the errors, the standard deviation of the error, and the root mean square error (RMSE) for each data set. Each of these statistical measures provides a measure of the quality of the TIN as compared to the original DEM data and can be used to help evaluate how well the TIN surface follows the DEM surface. These analyses were performed on the data from all 8 filtering thresholds for each of the three watershed areas. The following sections provide a brief summary of the statistical measures used and a summary of the statistical results for each data set.

### Maximum Error

The first statistical measure computed for each data set was the maximum error. This value is simply the largest of the elevation differences between the TIN surface and the original DEM grid, regardless of whether that elevation difference was positive or negative.

## Mean Error

The next statistical measure taken for each data set is the mean, which is used to describe the center or middle of a data set. For this application, if the actual error values derived from the TIN to DEM comparisons were used, then the average error value would be very near zero since the errors are positive and negative centered around zero. However, for the TIN to DEM comparison, it does not matter whether a TIN elevation is higher than the DEM or lower than the DEM. It is the difference between the two that is important. Thus, the absolute value of each error value was used in the computation of the mean, and the following equation for computing the mean was used:

$$\bar{x} = \frac{\sum_{i=1}^{n} |x_i|}{n} \qquad (11)$$

where $|x_i|$ represents the absolute value of the error at the $i^{th}$ grid point, and n is the total number of error values.

## Standard Deviation of Error

In addition to the maximum and mean value of any given data set, some measure of the extent to which the data is spread out or dispersed from the middle is useful in describing the data. The difference between any given error value in a sample and the mean error value for the entire sample is referred to as the deviation from the

mean. A standard approach in working with these deviations from the mean is the variance, which essentially measures the average of the squares of the deviations from the mean. The following equation is used to compute the variance, $s^2$:

$$s^2 = \frac{\sum_{i=1}^{n} ( x_i - \overline{x} )^2}{n - 1} \qquad (12)$$

where, $x_i$ is the error at the $i^{th}$ grid point, x bar is the mean error, and n is the number of error values in the sample. A more widely used measure of variation in a sample is the standard deviation. The standard deviation is defined as the square root of the variance, i.e., s. One of the advantages of the standard deviation is that it has the same units as the observations in the sample, which in this case is meters (Miller, Freund and Johnson 1990).

## Root Mean Square Error

One other statistical measure that is commonly used is the root mean square error (RMSE). To determine the RMSE, the error at each point is computed, squared, and summed for the entire set of data. The mean value of the squared errors is then taken, and the RMSE is the square root of that mean. This is given by the following equation:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} x_i^2}{n}} \qquad (13)$$

where, $x_i$ is the error (difference in elevation between the TIN and the DEM) at the $i^{th}$ DEM grid location, and n is the total number of error values in the data set.

Summary of Statistical Data

Table 2 provides a summary of the statistical measures described above for each data set used in this research. As would be expected, each of the statistical measures presented in Table 2 shows a general trend of degeneration as the filtering threshold increases and more data is filtered out of the DEM and fewer points are used to create the TIN. Also, the statistical measures indicate that the smallest errors occurred for the flatter topography of the Beaverdam Creek region and the errors increased as the roughness of the terrain increased for the Long Creek and Rock Creek areas.

There are a few exceptions to the trend of increasing error with increased filtering thresholds, all of which are in the maximum error data. For example, in the Rock Creek data, the maximum error at filter threshold 4 is greater than the maximum error at filter thresholds 5, 6, and 7. In such cases, there are typically one or two points out of the 20,000 to 30,000 points in the data set that have a very large error. An examination of the TINs reveals that these few points are always located along the edge of the TIN where some long, thin triangles may exist. Although the Delauney criterion eliminates almost all long, thin

Table 2. Summary of Statistical Error Data

**Summary of Statistical Error Data**

| ROCK CREEK | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Statistical Measure | Filtering Threshold / % of Data Filtered | | | | | | | |
| | 0 / 18.3 | 1 / 27.6 | 2 / 37.4 | 3 / 48.0 | 4 / 60.0 | 5 / 71.9 | 6 / 83.0 | 7 / 93.0 |
| Maximum Error, m | 27.5 | 30.0 | 36.2 | 64.8 | 117.9 | 64.7 | 67.9 | 111.2 |
| Mean Error, m | 0.26 | 0.49 | 0.75 | 1.09 | 1.75 | 2.40 | 4.24 | 8.98 |
| Std. Dev. of Error, m | 1.27 | 1.84 | 2.36 | 3.19 | 5.19 | 5.37 | 8.72 | 16.66 |
| RMSE, m | 1.25 | 1.78 | 2.25 | 3.02 | 4.89 | 4.86 | 7.81 | 14.30 |

| LONG CREEK | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Statistical Measure | Filtering Threshold / % of Data Filtered | | | | | | | |
| | 0 / 35.0 | 1 / 42.9 | 2 / 50.5 | 3 / 59.3 | 4 / 70.0 | 5 / 79.4 | 6 / 87.6 | 7 / 95.0 |
| Maximum Error, m | 4.0 | 6.3 | 11.6 | 19.0 | 20.2 | 9.0 | 20.3 | 20.3 |
| Mean Error, m | 0.02 | 0.05 | 0.08 | 0.12 | 0.18 | 0.26 | 0.42 | 0.94 |
| Std. Dev. of Error, m | 0.13 | 0.18 | 0.31 | 0.47 | 0.56 | 0.53 | 0.84 | 1.93 |
| RMSE, m | 0.13 | 0.18 | 0.30 | 0.45 | 0.53 | 0.47 | 0.74 | 1.72 |

| BEAVERDAM CREEK | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Statistical Measure | Filtering Threshold / % of Data Filtered | | | | | | | |
| | 0 / 38.2 | 1 / 44.6 | 2 / 50.7 | 3 / 61.4 | 4 / 75.6 | 5 / 83.1 | 6 / 89.8 | 7 / 96.2 |
| Maximum Error, m | 1.5 | 2.0 | 6.6 | 9.3 | 4.5 | 7.4 | 11.6 | 11.6 |
| Mean Error, m | 0.01 | 0.03 | 0.05 | 0.08 | 0.18 | 0.28 | 0.37 | 0.54 |
| Std. Dev. of Error, m | 0.09 | 0.13 | 0.20 | 0.28 | 0.36 | 0.49 | 0.65 | 0.95 |
| RMSE, m | 0.09 | 0.13 | 0.19 | 0.27 | 0.34 | 0.44 | 0.57 | 0.82 |

triangles, there are generally a few such triangles located

on the very edges of rectangular TINs. When a long, thin triangle exists, the vertices can be a significant distance from each other and the possibility of an extreme elevation difference is greatly increased. In addition to creating a larger than expected maximum error, these points may be of such an extreme as to distort the mean, standard deviation, and RMSE values. Although it was not pursued in this research, the statistical results obtained from this data filtering method could probably be improved simply by identifying these few extreme elevation differences along the edges of the TINs and adding a few additional vertices to the TIN at the proper locations so as to reduce or eliminate those very large error values. Not only would this reduce the maximum error, but it would also ensure that the mean and standard deviation values are not unduly influenced by a few extreme values. Figure 35 shows an example of how these long, thin triangles can exist along the edge of a rectangular tin.

The information provided in this chapter is useful for numerically comparing the quality of the TINs created and to illustrate how that quality changes as the filtering threshold changes. However, such statistical information does little to tell a hydrologist how watershed delineation and basin geometry is effected by the various filtering thresholds. The following chapters consider those impacts.

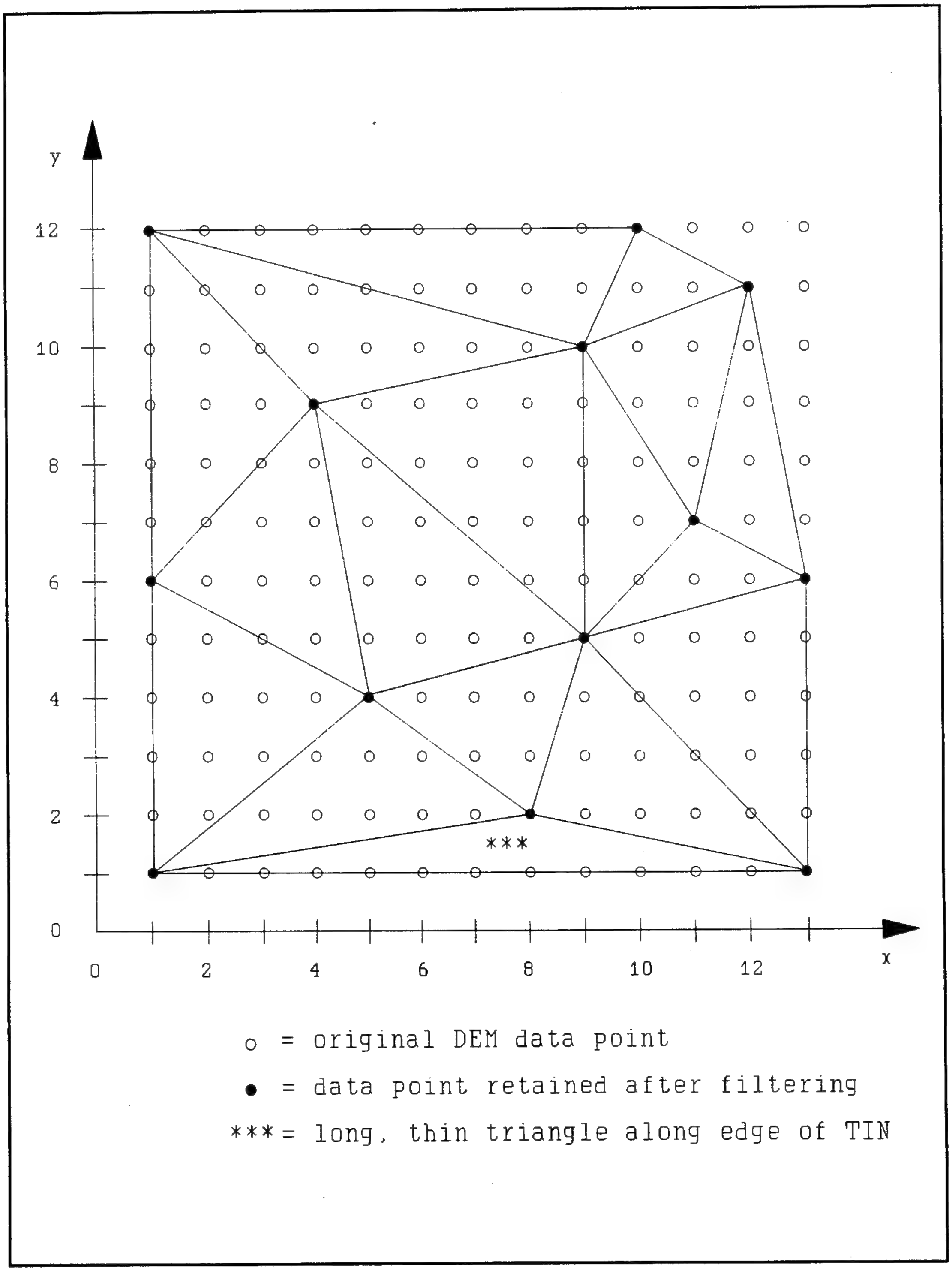



**Figure 35. Long, Thin Triangle along TIN Edge**

# CHAPTER VIII

# WATERSHED GEOMETRIC PARAMETERS

As discussed in detail in Chapter V, the DEM elevation data filtering technique being analyzed in this research allows the user to choose from a range of threshold values. Each of those threshold values will result in a different number of surface specific points being selected, a different TIN, and a different level of accuracy relative to the original DEM. The previous chapter discussed some of the statistical implications of the data filtering process. However, the real interest in such data filtering is its effect on modeling the hydrology of a watershed. Thus, the specific focus of this research is to determine the effect of the filtering when using the resultant TIN to delineate a watershed boundary, compute the geometry associated with that watershed, and finally use those watershed geometric parameters in a lumped parameter hydrologic model, such as HEC-1, to model the watershed. The following sections detail the results of watershed delineation and computation of watershed geometric parameters using the TINs created from various levels of data filtering, and the following chapter summarizes the HEC-1 model results obtained using those geometric data.

## Watershed Delineation

Although the process of delineating a watershed, as outlined earlier in this report, appears to be fairly

straightforward, it can be somewhat time consuming. This is particularly true when using a TIN created from data filtered at a very low threshold level. Those TINs tend to have large, relatively flat regions that were not removed during the filtering process, and a certain amount of additional processing is required in order to complete the basin delineation. The TINs created from data filtered at higher threshold levels have very few flat regions, and can be delineated much more quickly. The time saved by using more highly filtered data is one of the primary reasons why DEM is filtering is advantageous for TIN creation.

For the purposes of this research, the watersheds for each of the three regions were not delineated for all eight of the filtering thresholds. In addition to performing watershed delineation using the unfiltered DEM grid, four levels of filtering were chosen, threshold values 1, 3, 5, and 7. This range of data sets provide ample information on the effect of data filtering on watershed delineation.

Using the TINs for each of these filtering thresholds in each of the three geographic regions, the watershed within each region was delineated. This was accomplished by first identifying the fixed location on each TIN of the basin outlet. For example, in the Rock Creek area the basin outlet was located at UTM coordinates 452675,3732143 and that location was identified as the basin outlet for each of the TINs for the Rock Creek area. Using that basin

outlet point as the terminus, the watershed delineation technique outlined in Chapter III was used to determine the set of triangles which contributed flow to that point. That set of triangles was considered to be the watershed area for that outlet. For each of the three topographic regions, this process was performed on a TIN generated from the original DEM grid with no filtering, as well as TINs generated from DEM data filtered at thresholds 1, 3, 5 and 7. This resulted in 5 watershed delineations for each of the three regions, or a total of 15 watersheds.

## Watershed Geometry

Using the watersheds defined by the above process, the basin geometric parameters detailed in Chapter III were computed for each watershed. Thus, the area, average slope, and maximum flow distance were computed for each of the data filtering thresholds in each geographic region. The basin geometric parameters computed for the TIN which was generated from the original DEM data with no filtering were assumed to be the most correct values. These geometric parameters were used as the basis for comparison of the geometric values that were then computed for the TINs generated from the filtered DEM data at the 1, 3, 5, and 7 threshold levels. The results of this comparison provide insight into the impacts of the data filtering process on the computation of these parameters. As fewer and fewer of the DEM data points are used to create a TIN,

one would expect that the accuracy of the basin geometric parameters would deteriorate. A summary of the basin geometric parameters computed for the 15 cases in this research is presented in Table 3.

In general, the variation in the geometric parameters presented in Table 3 tends to increase as fewer and fewer of the DEM points are used in the TIN. However, that trend is not as clear as expected. For example, the average slope for the Beaverdam Creek data at the filter threshold of 7 is only 1.79% greater than the slope from the original DEM TIN, while the average slope for that watershed at the filter thresholds of 1, 3, and 5 are much greater at 7.14%, 5.36%, and 7.14% respectively. In most of the data, however, the trend is towards a greater variation as fewer data points are used. Obviously, this research was conducted on a limited number of cases and some anomalous results are to be expected. If this process were to be repeated on a large number of watersheds in various types of terrain, the overall trend would more clearly show that the geometric parameters at the higher filtering thresholds vary more from the original values that do those at the lower thresholds.

Table 3. Summary of Watershed Geometric Parameters

| Watershed Geometric Parameters |
| --- |

| ROCK CREEK | | | | | | |
|---|---|---|---|---|---|---|
| Geometric Parameter | | Filter Threshold / % of Data Filtered | | | | |
| | | None / 0 | 1 / 27.6 | 3 / 48.0 | 5 / 71.9 | 7 / 93.0 |
| Area (km2) | Value | 65.71 | 65.10 | 65.31 | 64.12 | 69.07 |
| | % Diff. | - | -0.93 | -0.61 | -2.42 | +5.11 |
| Average Slope | Value | .2040 | .2040 | .2038 | .2091 | .2044 |
| | % Diff. | - | 0.00 | -0.10 | +2.50 | +0.20 |
| Max. Flow Distance (m) | Value | 17023 | 16410 | 16770 | 16594 | 15861 |
| | % Diff. | - | -3.61 | -1.495 | -2.52 | -6.833 |

| LONG CREEK | | | | | | |
|---|---|---|---|---|---|---|
| Geometric Parameter | | Filter Threshold / % of Data Filtered | | | | |
| | | None / 0 | 1 / 42.9 | 3 / 59.3 | 5 / 79.4 | 7 / 95.0 |
| Area (km2) | Value | 81.23 | 81.71 | 81.42 | 80.40 | 80.78 |
| | % Diff. | - | +0.59 | +0.23 | -1.02 | -0.55 |
| Average Slope | Value | .0121 | .0122 | .0123 | .0128 | .0133 |
| | % Diff. | - | +0.83 | +1.65 | +5.79 | +9.92 |
| Max. Flow Distance (m) | Value | 16415 | 16727 | 16354 | 16385 | 15373 |
| | % Diff. | - | +1.88 | -0.39 | -0.20 | -6.37 |

| BEAVERDAM CREEK | | | | | | |
|---|---|---|---|---|---|---|
| Geometric Parameter | | Filter Threshold / % of Data Filtered | | | | |
| | | None / 0 | 1 / 44.6 | 3 / 61.4 | 5 / 83.1 | 7 / 96.2 |
| Area (km2) | Value | 43.21 | 42.76 | 41.53 | 40.90 | 44.85 |
| | % Diff. | - | -1.25 | -3.89 | -5.35 | +3.80 |
| Average Slope | Value | .0056 | .0060 | .0059 | .0060 | .0057 |
| | % Diff. | - | +7.14 | +5.36 | +7.14 | +1.79 |
| Max. Flow Distance (m) | Value | 12407 | 11634 | 11098 | 11193 | 10728 |
| | % Diff. | - | -6.23 | -10.55 | -9.79 | -13.53 |

These geometric parameters are primarily of interest

because of their potential use in hydrologic modeling. In particular for this research, they represent input variables to the lumped parameter HEC-1 model. Thus, the following chapter will focus on the major thrust of this research, i.e., how the various elevation data filtering thresholds impact lumped parameter hydrologic model results.

# CHAPTER IX

# HYDROLOGIC MODELING

As noted in the previous two chapters, the statistical measures and basin geometric parameters for the various level of data filtering are informative, but they still do not provide specific guidance on the impacts of this data filtering process on the actual output from a hydrologic model. To provide that information, the basin geometric parameters from Chapter VIII were used to generate input parameters for the HEC-1 runoff model. Basic information required to run HEC-1 for a single basin are rainfall data, loss rate parameters, basin area, and unit hydrograph parameters. To obtain a comparison of results only dependent on elevation data filtering, all input data not dependent on the basin geometry was taken to be constant for all model runs.

The intent of this effort was to compare the runoff hydrographs that were computed for each of the three watersheds by using the data sets for each watershed based on the 5 levels of filtering, i.e., no filtering and filtering at threshold levels 1, 3, 5, and 7. For each watershed the base case was the hydrograph computed using the basin geometric parameters from the TIN generated with

the unfiltered DEM data. Hydrographs resulting from the filtered data at the threshold of 1, 3, 5, and 7 were then compared to that. For each case, the rainfall was kept constant and the same loss rate parameters were used. The input variables that did differ were those based upon the basin geometry, specifically the basin area and the unit hydrograph parameter, where the unit hydrograph parameter was computed by using the basin average slope and maximum flow distance values as detailed later in this chapter.

## The HEC-1 Model

The lumped parameter model used in this research was the HEC-1 model developed by the U.S. Army Corps of Engineers, Hydrologic Engineering Center. HEC-1 simulates the surface runoff response of a watershed to a given precipitation event. The watershed is typically subdivided into sub-basins, each of which represents a surface runoff entity with the sub-basins connected by stream channels. Thus, the runoff from each sub-basin is computed, combined with runoff from other sub-basins, and routed through the channels to derive a hydrograph at the desired location. A basic assumption in HEC-1 is that the hydrologic processes in each sub-basin can be represented by model parameters which reflect average conditions within the sub-basin. HEC-1 simulations are limited to single storm events, as there is no provision for soil moisture recovery between storms (U.S. Army Corps of Engineers 1990).

The HEC-1 user's manual provides detailed information on the structure and requirements of HEC-1 input files for various applications of the model. The specific requirements of HEC-1 will not be detailed in this report, but the options used in modeling the various watersheds in this study will be briefly discussed. The following sections detail the major components of the HEC-1 input files created for this study.

## Model Input

The input for the HEC-1 model runs that were made for this research can be divided into two types of input, one type may be called "constant parameters" and the other may be called "geometric based parameters." The constant parameters are those that are not based on the basin geometry, and thus were held constant for every watershed. These include the rainfall and loss rate parameters. The geometric based parameters are those based on the geometry of the basin and were varied for each watershed, depending upon the basin geometry derived from the TIN for each filtering threshold. These geometric based parameters include the basin area and unit hydrograph parameter. The following sections summarize these types of input data.

### Rainfall

There are several methods of using rainfall input in the HEC-1 model. These include using a basin average, gage

data, or various hypothetical storm simulations. For the purpose of this study, the main requirements for the rainfall input were that it be the same for each watershed and of sufficient magnitude to generate easily observable runoff from each watershed. An average rainfall of 50.8 mm (equivalent to 2 inches) was chosen as a basin average precipitation to satisfy these conditions.

The other component of the rainfall input is the time distribution of that rainfall, or the rainfall hyetograph. Again, any distribution of rainfall would serve the purpose of this study, as long as the same distribution was used for each case. Since the SCS Type II 24-hour standard distribution is applicable to the geographic regions for all three of the study watersheds presented here, it was used for this study. Table 4 shows this particular distribution (Chow, Maidment, and Mays 1988).

## Loss Rate Method

In any hydrologic model, some means of estimating the amount of rainfall that becomes runoff must be used. During a rainfall event, some portion of the rainfall is lost to infiltration, evapotranspiration, interception, and other minor losses. Some portion of the losses are termed "initial abstractions" in that they must be satisfied prior to any runoff occurring. Additional losses continue to occur after runoff has commenced. To model this effect, a loss rate parameter must be established to account for the

Table 4. SCS Type II 24-Hour Distribution

| SCS Type II 24-Hour Distribution | |
|---|---|
| Hour | $P_t$ / $P_{24}$ |
| 0.0 | 0.0 |
| 2.0 | 0.022 |
| 4.0 | 0.048 |
| 6.0 | 0.080 |
| 7.0 | 0.098 |
| 8.0 | 0.120 |
| 8.5 | 0.133 |
| 9.0 | 0.147 |
| 9.5 | 0.163 |
| 9.75 | 0.172 |
| 10.0 | 0.181 |
| 10.5 | 0.204 |
| 11.0 | 0.235 |
| 11.5 | 0.283 |
| 11.75 | 0.357 |
| 12.0 | 0.663 |
| 12.5 | 0.735 |
| 13.0 | 0.772 |
| 13.5 | 0.799 |
| 14.0 | 0.820 |
| 16.0 | 0.880 |
| 20.0 | 0.952 |
| 24.0 | 1.000 |

losses. After these losses have been estimated, the remaining rainfall, referred to as excess rainfall, becomes runoff.

The loss rate method chosen for this study was the SCS Curve Number method developed by the Soil Conservation Service. The SCS Curve Number method was developed by studying the rainfall-runoff relationships from many small experimental watersheds. This method requires the assignment of a single parameter referred to as the curve number or CN. The CN is a dimensionless parameter with valid values between 0 and 100. Higher values of the CN represent fewer losses and a higher amount of runoff, while lower values of CN represent higher loss rates and thus a smaller amount of excess rainfall which can become runoff. For example, a totally impervious area would be given a CN of 100 while natural surfaces such as pastures, meadows, woodlands, or cultivated land would have a CN less than 100 since some amount of losses will occur in these areas.

Derivation of the CN is based on the principal that the depth of excess precipitation or runoff, $P_e$, will always be less than or equal to the depth of precipitation, P. This excess precipitation, $P_e$, is the amount remaining after initial abstractions, $I_a$, and continuing abstractions, $F_a$. These continuing abstractions are those that continue after runoff begins, and represent the additional depth of water retained in the watershed. $F_a$ will be less than or equal to

the potential maximum retention, S. Figure 36 illustrates these relationships.

The potential runoff from a basin is defined as the total amount of precipitation remaining after the initial abstractions are satisfied, or $P-I_a$. The basic hypothesis of the SCS method is that the ratios between the two actual and potential quantities described here are equal, i.e.,

$$\frac{F_a}{S} = \frac{P_e}{P - I_a} \tag{14}$$

From the continuity principle for the values depicted in Figure 36, it can be seen that

$$P = P_e + I_a + F_a \tag{15}$$

Combining these two equations and solving for $P_e$ gives

$$P_e = \frac{(P - I_a)^2}{P - I_a + S} \tag{16}$$

This equation is used for computing the depth of excess rainfall, or direct runoff, from a storm by the SCS method.

By study of results from a variety of watersheds, the following empirical relationship between $I_a$ and S was developed:

$$I_a = 0.2\ S \tag{17}$$

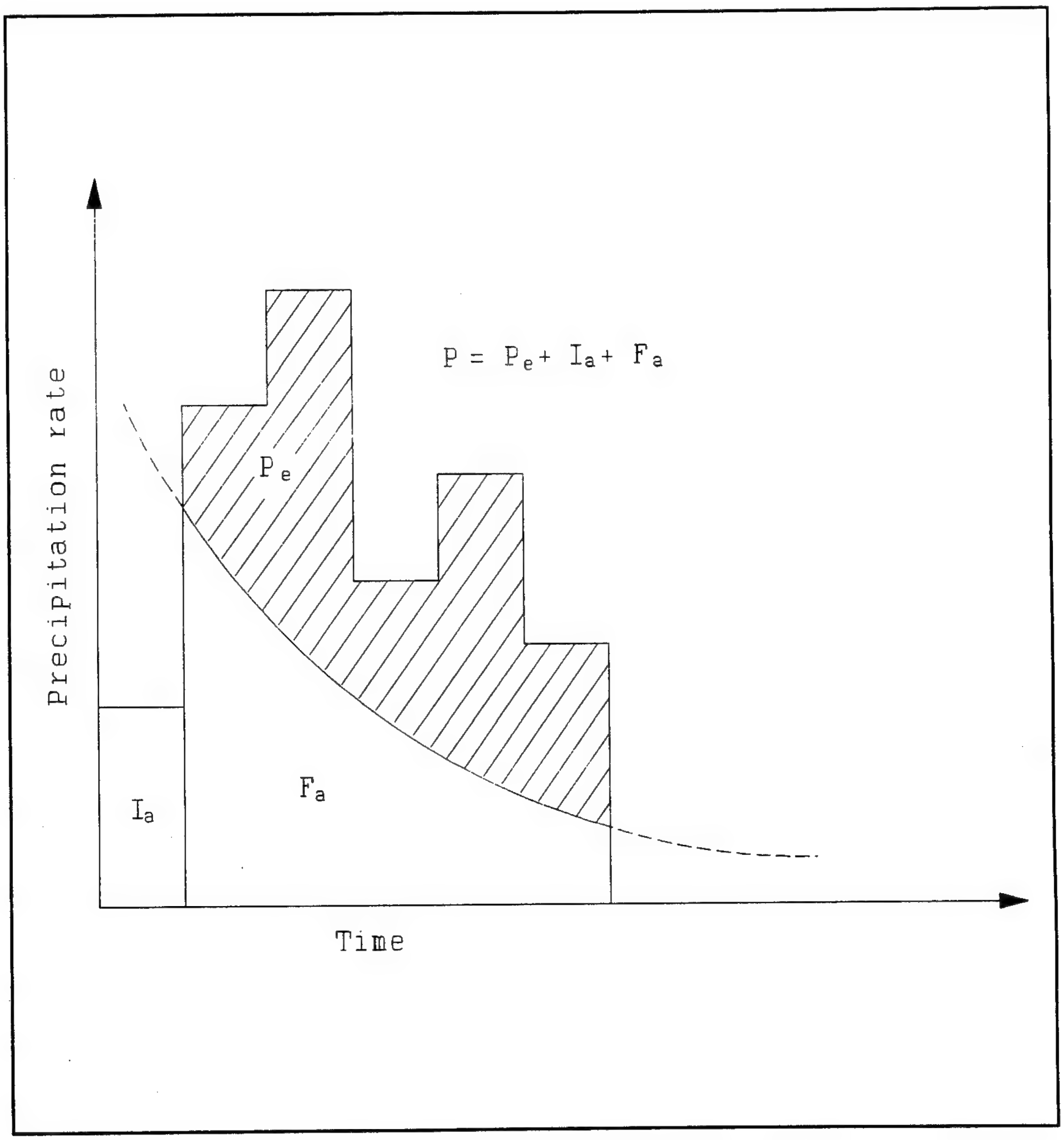



Figure 36. Excess Precipitation and Losses

Using this relationship in equation 16 yields

$$P_e = \frac{(P - 0.2\ S)^2}{P + 0.8\ S} \tag{18}$$

Plotting data for P and $P_e$ from many watersheds, the SCS established a family of curves. These curves were standardized by using the dimensionless curve number, CN, which is related to S by

$$S = \frac{1000}{CN} - 10 \tag{19}$$

with S in inches. Thus, for a cumulative rainfall, P, and runoff curve number, CN, the cumulative direct runoff, $P_e$, can be determined (Chow, Maidment, and Mays 1988).

The SCS Curve Number method of computing losses was chosen for this research for several reasons. The first is that the unit hydrograph method chosen, which is discussed in the following section, has the CN as a variable. Thus, a CN must be assigned to each watershed to compute the unit hydrograph as described below. Secondly, the SCS Curve Number method is widely used, and only requires one parameter to define the losses for a particular basin. As with the other constant parameters of this study, the actual value of CN chosen was not as important as the fact that the CN be kept the same for each case. A CN of 80 was chosen for each watershed in this study.

## Unit Hydrograph Method

The unit hydrograph was first proposed by Sherman in 1932 and was originally called simply the unit-graph. The unit hydrograph for a watershed is defined as a direct runoff hydrograph resulting from one unit of excess rainfall generated uniformly over the drainage area at a constant rate for an effective duration (Chow, Maidment, and Mays 1988). There are several methods for generating a synthetic unit hydrograph for a basin or watershed, many of which employ some of the geometric parameters of the watershed in determining the unit hydrograph parameters.

The unit hydrograph chosen for this study is the SCS Dimensionless Unit Hydrograph. The SCS Dimensionless Unit Hydrograph was derived from a large number of natural unit hydrographs from watersheds varying widely in size and geographical locations (U.S. Department of Agriculture 1985). This dimensionless hydrograph has its ordinate values expressed in the dimensionless ratio $q/q_p$, where q is the discharge at time t and $q_p$ is the peak discharge. The abscissa values are expressed by the dimensionless quantity $t/T_p$, where t is the time and $T_p$ is the time from the beginning of the rise to the peak. Figure 37 provides a plot of the SCS Dimensionless Unit Hydrograph.

In the HEC-1 model, the SCS unit hydrograph is defined by a single parameter called TLAG, which is defined as the lag in hours between the center of mass of rainfall excess

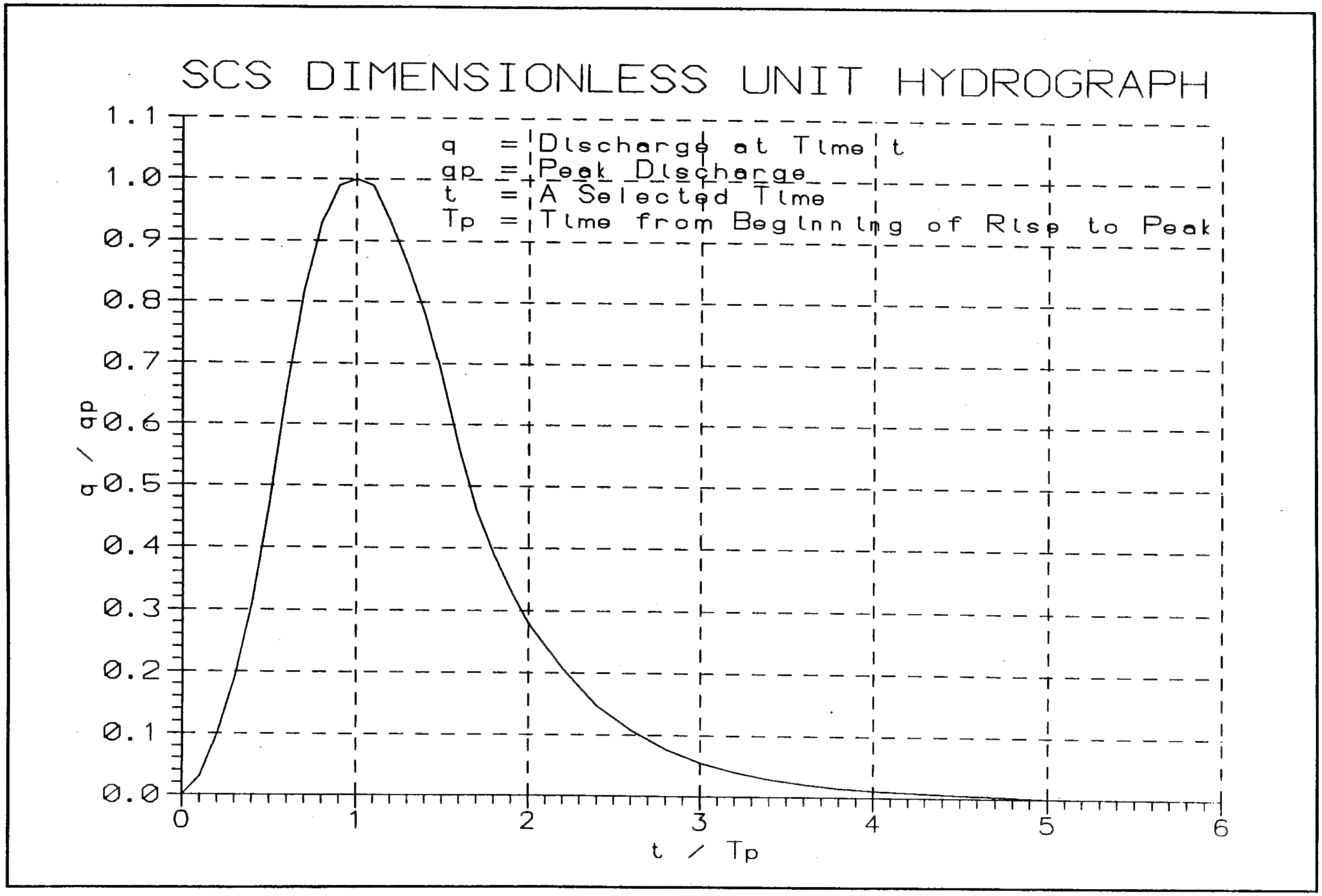



Figure 37. SCS Dimensionless Unit Hydrograph

and the peak of the unit hydrograph. In HEC-1, the SCS Dimensionless Unit Hydrograph parameters $T_p$ and $q_p$ are computed as follows:

$$T_p = 0.5 * \Delta t + TLAG \tag{20}$$

$$q_p = 484 * AREA / T_p \tag{21}$$

In the above equations, $\Delta t$ is taken as the HEC-1 computational time interval and AREA is the area of the watershed or subbasin over which the unit hydrograph is to be applied. The value for AREA in this research was taken as the area determined from the TINs, as discussed in the following section. The only remaining value needed then is that of TLAG, and then the unit hydrograph can be determined from the $q/q_p$ and $t/T_p$ ratios shown in Figure 37.

The computation of TLAG can be done in several manners. The SCS (U.S. Department of Agriculture 1985) presents an equation which is based upon some of the watershed geometric parameters which can be computed from a TIN. That equation is as follows:

$$TLAG = \frac{L^{0.8} * (S + 1)^{0.7}}{1900 * Y^{0.5}} \tag{22}$$

where, L = maximum flow distance in the watershed (ft)

S = (1000/CN) - 10

Y = average watershed slope (%)

This provides a single equation which incorporates two of the watershed geometric parameters that can be calculated from a TIN model, i.e., the maximum flow distance and the average watershed slope. This is the approach used in this research to compute TLAG for input into HEC-1. The values for maximum flow distance and average watershed slope computed from the TINs for various filtering threshold levels, as presented in Table 3, were used to determine a TLAG value for each watershed.

## Watershed Area

As discussed in detail in Chapter III, computation of the area of a watershed from a TIN model is accomplished through a summation of the areas of all of the triangles which belong to that watershed. For the HEC-1 input files in this research, the area of each watershed was taken as the area computed from the TIN model. Since the UTM coordinate system, which was used for the DEM data, is in metric units, the areas computed by this process were given in square kilometers ($km^2$).

## Input Data Summary

Although only one HEC-1 run for each data set is presented here, simulations were made using different rainfall and CN values. However, the results from those efforts added no additional insight. Although the peak flow and time to peak changed as those parameters were

changed, the results of the five test cases for each watershed remained essentially the same relative to each other. Thus, the conclusions derived from results presented here for a uniform rainfall of 50.8 mm and a CN of 80 are the same conclusions as those that would be derived from the results of other rainfall and CN values. Table 5 presents a summary of the HEC-1 input data for each of the 15 test cases presented here, with the same 50.8 mm of rainfall and CN of 80 used in each case.

## Model Results

The HEC-1 input data summarized in Table 5 was used to generate input files for HEC-1 for each of the 15 cases presented. With those input files, the HEC-1 model was run for each case, and the resulting hydrographs were computed. The hydrographs from these HEC-1 model runs are presented in Figures 38, 39, and 40, for Rock Creek, Long Creek, and Beaverdam Creek respectively. The hydrographs resulting from the unfiltered data are shown with a solid line, while those for the filtering thresholds of 1, 3, 5, and 7 are shown by varying degrees of dashed lines as indicated in the legends of those Figures.

Although the hydrographs presented in Figures 38, 39, and 40 are visually informative, a summary of the data in tabular format also provides a useful means of comparing the data. Table 6 provides a summary of the peak flow and time of peak flow for each of the test cases. As before,

Table 5. Summary of HEC-1 Input Data

| Summary of HEC-1 Input Data | | | | | |
|---|---|---|---|---|---|
| ROCK CREEK | | | | | |
| Input Variable | Filter Threshold / % of Data Filtered | | | | |
| | None / 0 | 1 / 27.6 | 3 / 48.0 | 5 / 71.9 | 7 / 93.0 |
| Area ($km^2$) | 65.71 | 65.10 | 65.31 | 64.12 | 69.07 |
| TLAG (hr) | 1.76 | 1.71 | 1.74 | 1.70 | 1.66 |
| LONG CREEK | | | | | |
| Input Variable | Filter Threshold / % of Data Filtered | | | | |
| | None / 0 | 1 / 42.9 | 3 / 59.3 | 5 / 79.4 | 7 / 95.0 |
| Area ($km^2$) | 81.23 | 81.71 | 81.42 | 80.40 | 80.78 |
| TLAG (hr) | 7.01 | 7.09 | 6.93 | 6.80 | 6.34 |
| BEAVERDAM CREEK | | | | | |
| Input Variable | Filter Threshold / % of Data Filtered | | | | |
| | None / 0 | 1 / 44.6 | 3 / 61.4 | 5 / 83.1 | 7 / 96.2 |
| Area ($km^2$) | 43.21 | 42.67 | 41.53 | 40.90 | 44.85 |
| TLAG (hr) | 8.24 | 7.56 | 7.34 | 7.33 | 7.27 |

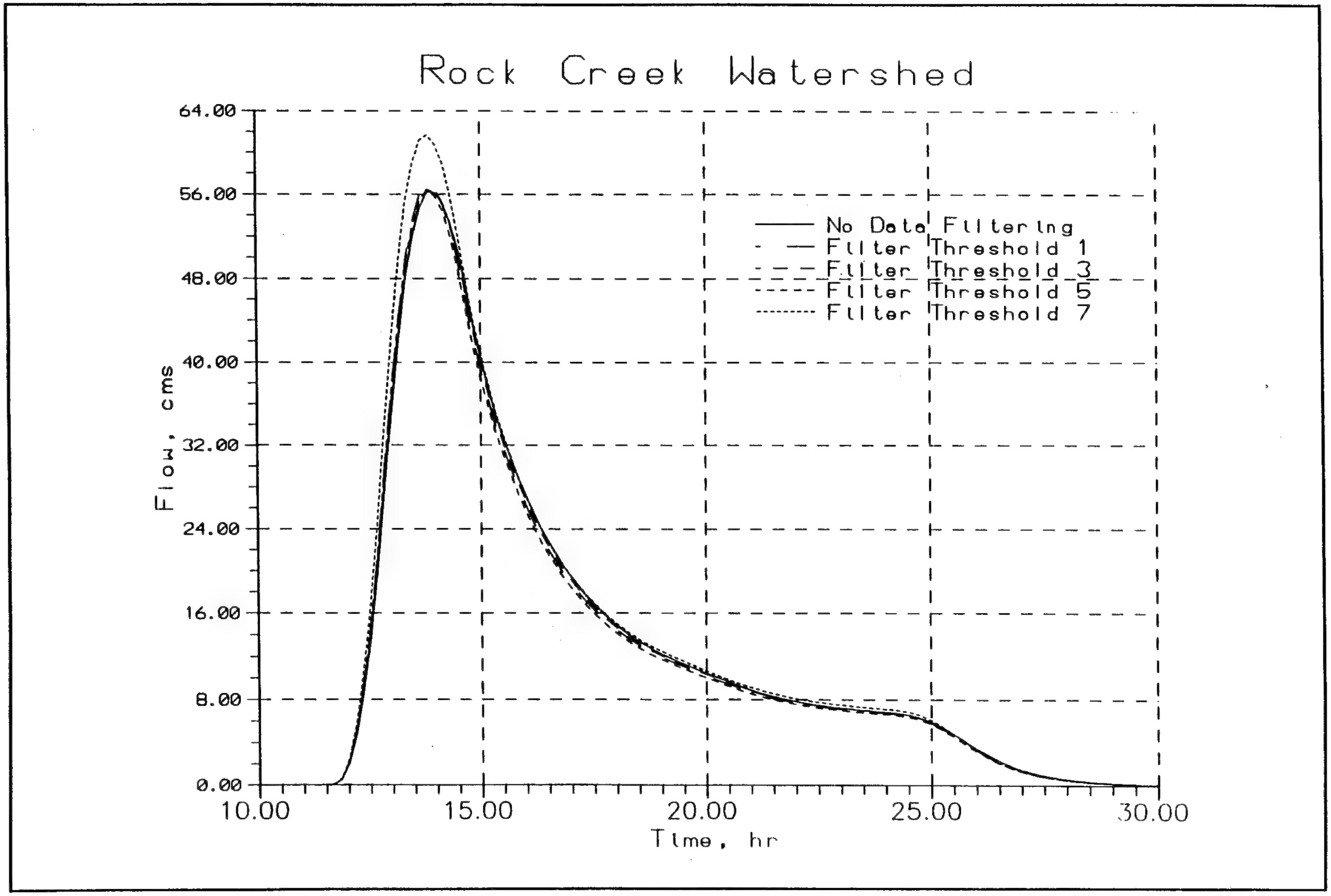



Figure 38. Rock Creek Model Results

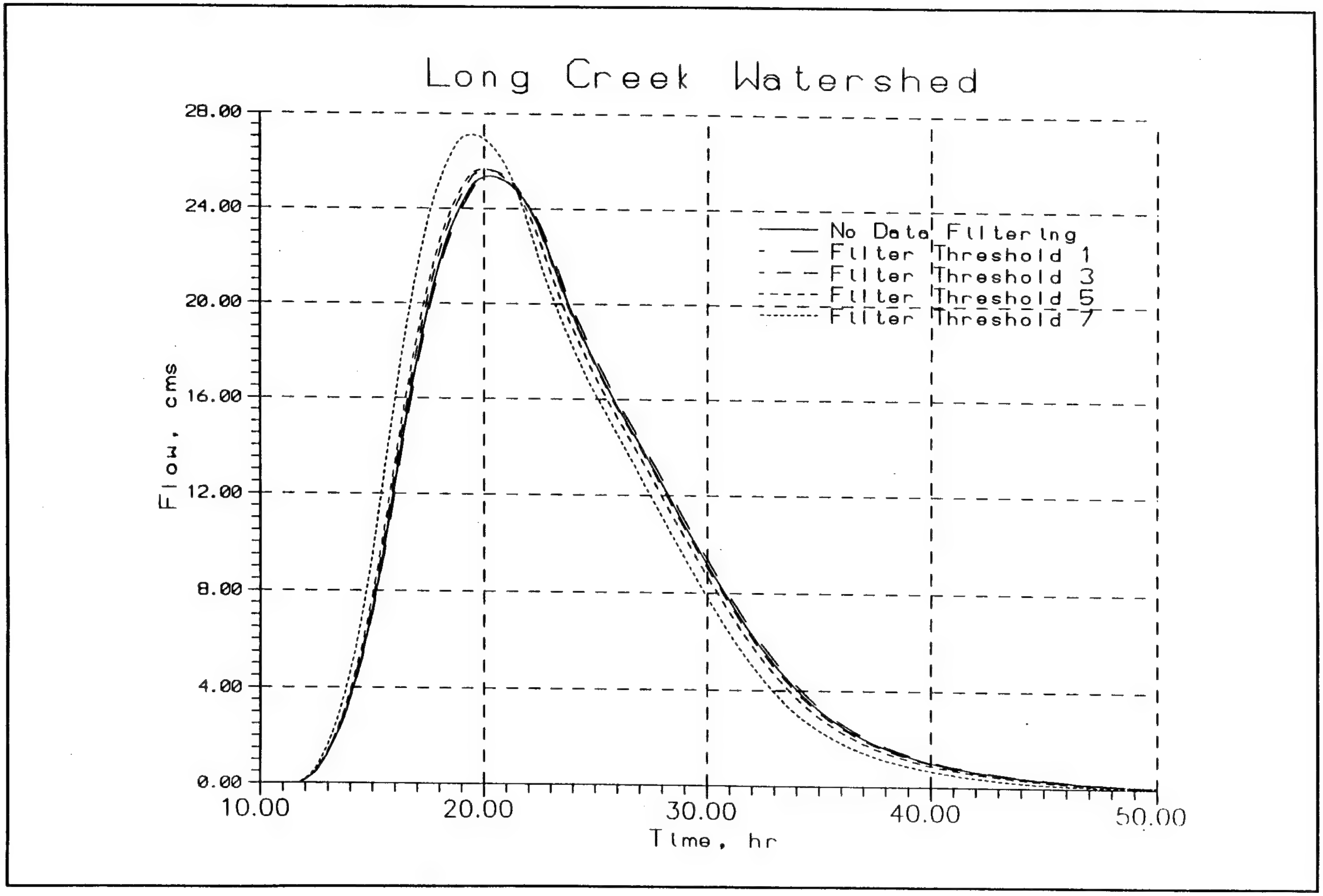



Figure 39. Long Creek Model Results

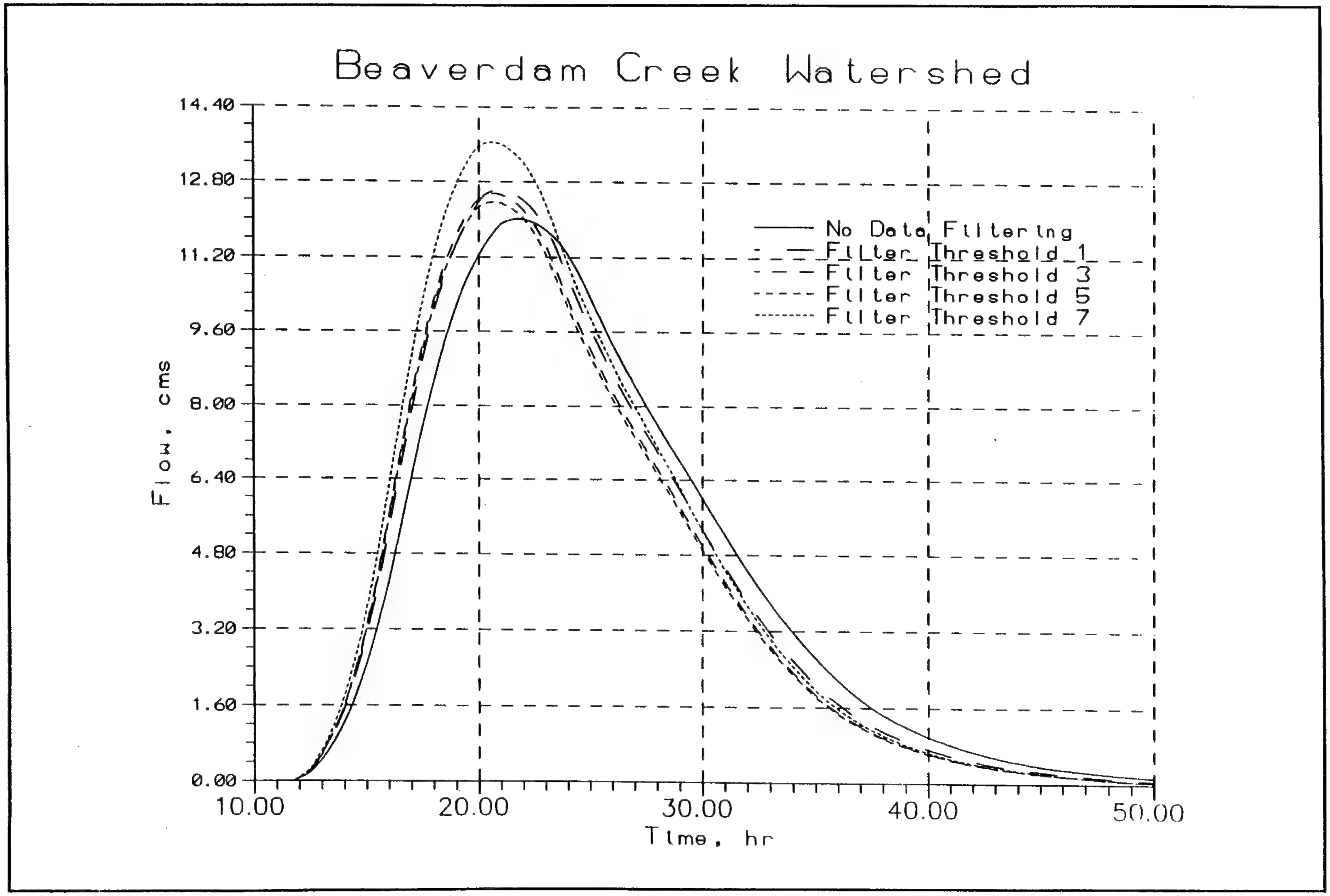



Figure 40. Beaverdam Creek Model Results

Table 6. Summary of Runoff Hydrograph Results

## Summary of Runoff Hydrograph Results

### ROCK CREEK

| Runoff Parameter | | Filter Threshold / % of Data Filtered | | | | |
|---|---|---|---|---|---|---|
| | | None/ 0 | 1 / 27.6 | 3 / 48.0 | 5 / 71.9 | 7 / 93.0 |
| Peak Flow | Value ($m^3/s$) | 56.20 | 56.99 | 56.45 | 56.36 | 61.69 |
| | % Diff. | - | +1.41 | +0.44 | +0.28 | +9.77 |
| Time to Peak | Value (hr) | 13.83 | 13.83 | 13.83 | 13.83 | 13.83 |
| | % Diff. | - | 0.00 | 0.00 | 0.00 | 0.00 |

### LONG CREEK

| Runoff Parameter | | Filter Threshold / % of Data Filtered | | | | |
|---|---|---|---|---|---|---|
| | | None/ 0 | 1 / 42.9 | 3 / 59.3 | 5 / 79.4 | 7 / 95.0 |
| Peak Flow | Value ($m^3/s$) | 25.37 | 25.31 | 25.64 | 25.67 | 27.12 |
| | % Diff. | - | -0.24 | +1.06 | +1.18 | +6.90 |
| Time to Peak | Value (hr) | 20.25 | 20.25 | 20.25 | 20.00 | 19.50 |
| | % Diff. | - | 0.00 | 0.00 | -1.23 | -3.70 |

### BEAVERDAM CREEK

| Runoff Parameter | | Filter Threshold / % of Data Filtered | | | | |
|---|---|---|---|---|---|---|
| | | None/ 0 | 1 / 44.6 | 3 / 61.4 | 5 / 83.1 | 7 / 96.2 |
| Peak Flow | Value ($m^3/s$) | 12.01 | 12.62 | 12.55 | 12.37 | 13.64 |
| | % Diff. | - | +5.08 | +4.50 | +3.00 | +13.57 |
| Time to Peak | Value (hr) | 21.75 | 21.00 | 20.75 | 20.75 | 20.75 |
| | % Diff. | - | -3.45 | -4.60 | -4.60 | -4.60 |

the hydrograph data derived from the unfiltered original DEM data was used as a base case, and the results of all other cases were compared to that, with the resultant percent differences also presented in Table 6. It can be seen from the data in Table 6 that the percent change in the peak flow for each of the watershed increases much more dramatically at the filter threshold of 7. In fact, for the Rock Creek and Beaverdam Creek data, the percent difference in the peak flow value actually decreases as it moves from the filter threshold of 1 through the threshold of 3 to the threshold of 5. The peak flow for the Long Creek data has a basically insignificant change over these same thresholds. The real increase comes at the level 7 filter threshold, and this is consistent for all three of the different terrains represented here. With only one exception in all of these data sets, the peak flow is always greater when using filtered data than it is for the unfiltered DEM data. In all cases where there is a change in the time to peak, the time to peak always occurs earlier for the filtered data cases than for the unfiltered data cases.

The trends in the peak flow data presented in Table 6 can be more easily seen when the percent difference in the peak flow is plotted against the percentage of data filtered from the original DEM grid, as shown in Figures 41, 42, and 43. As can be seen from these plots, the

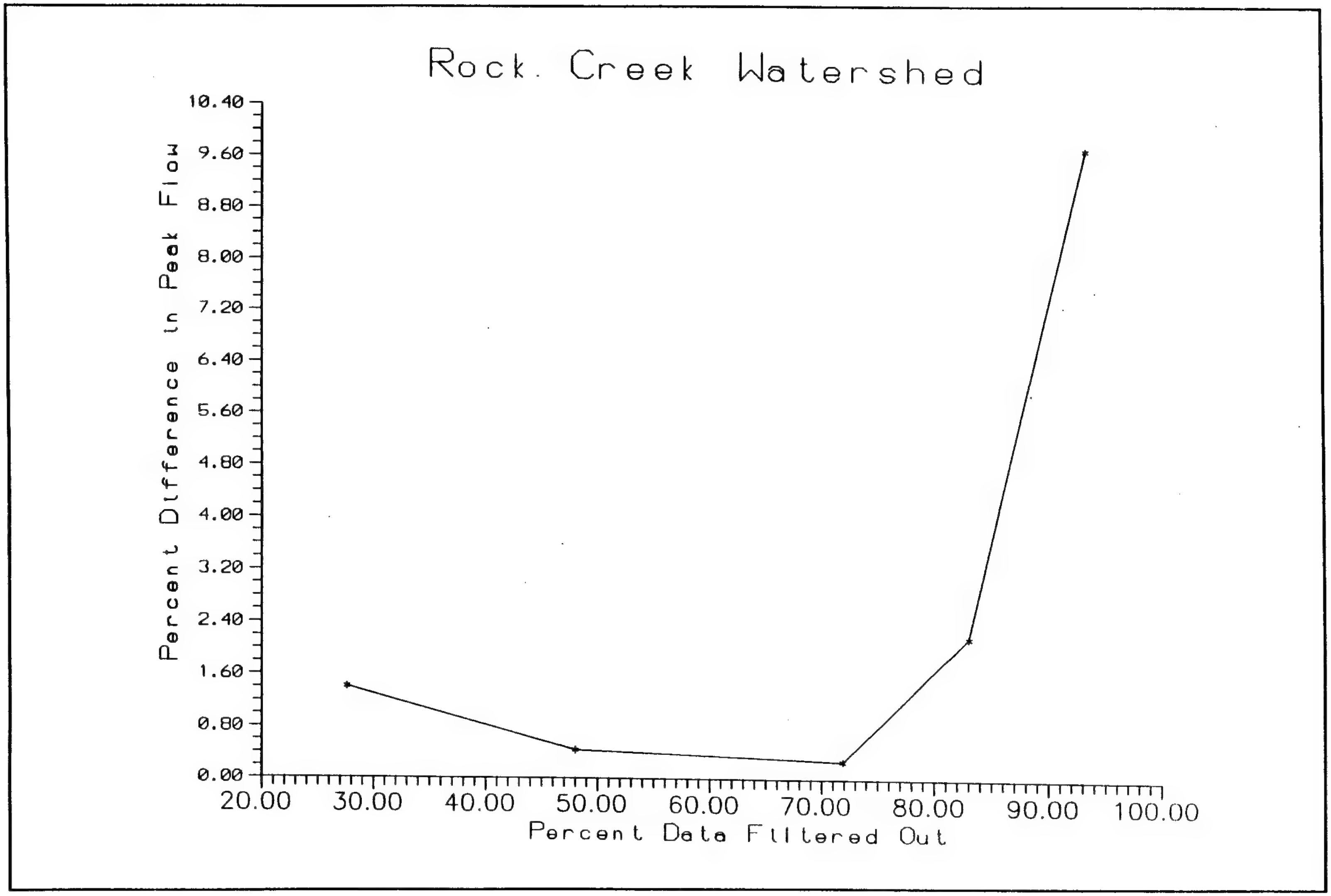



Figure 41. Peak Flow Changes vs. Data Filtered for Rock Creek

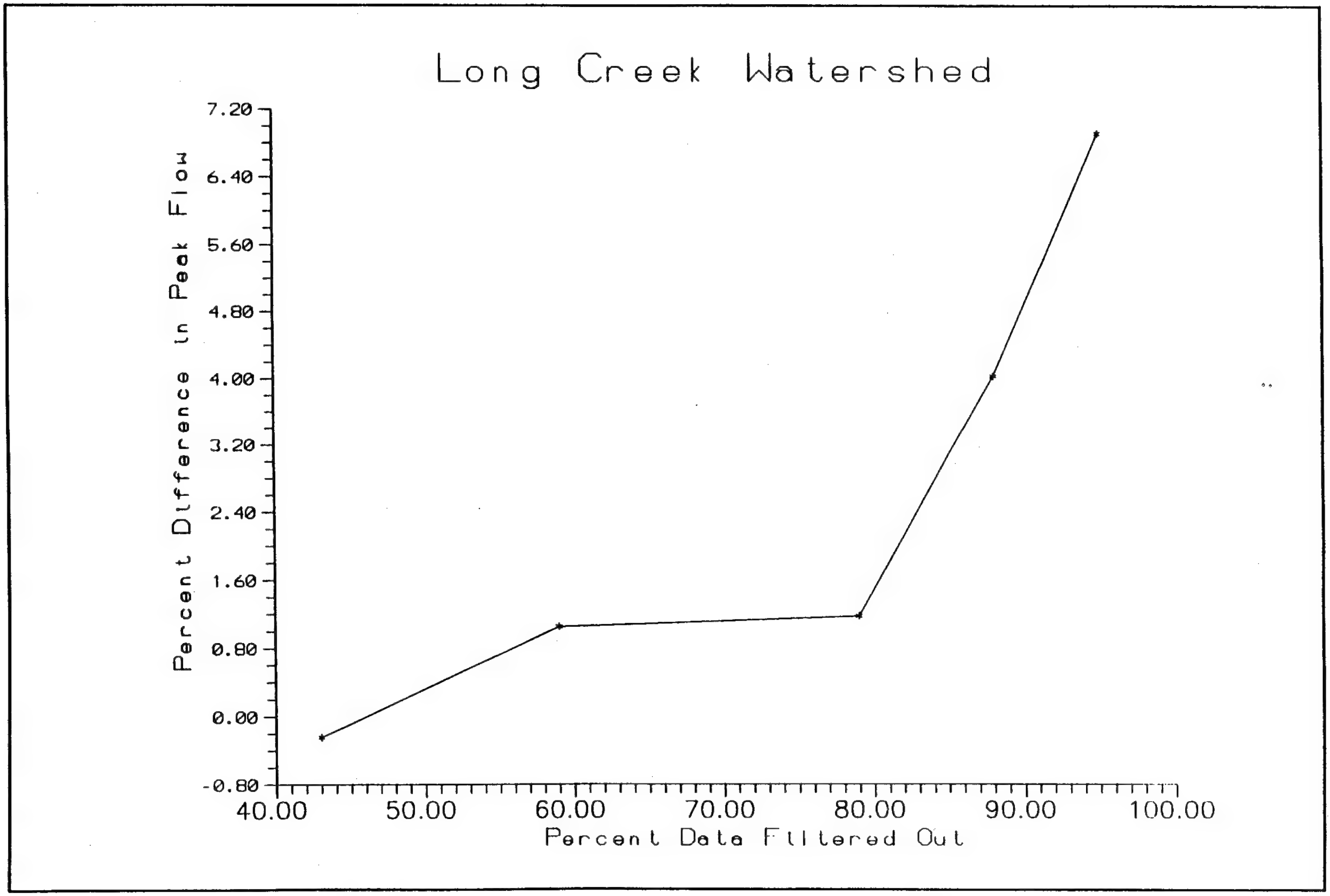



Figure 42. Peak Flow Changes vs. Data Filtered for Long Creek

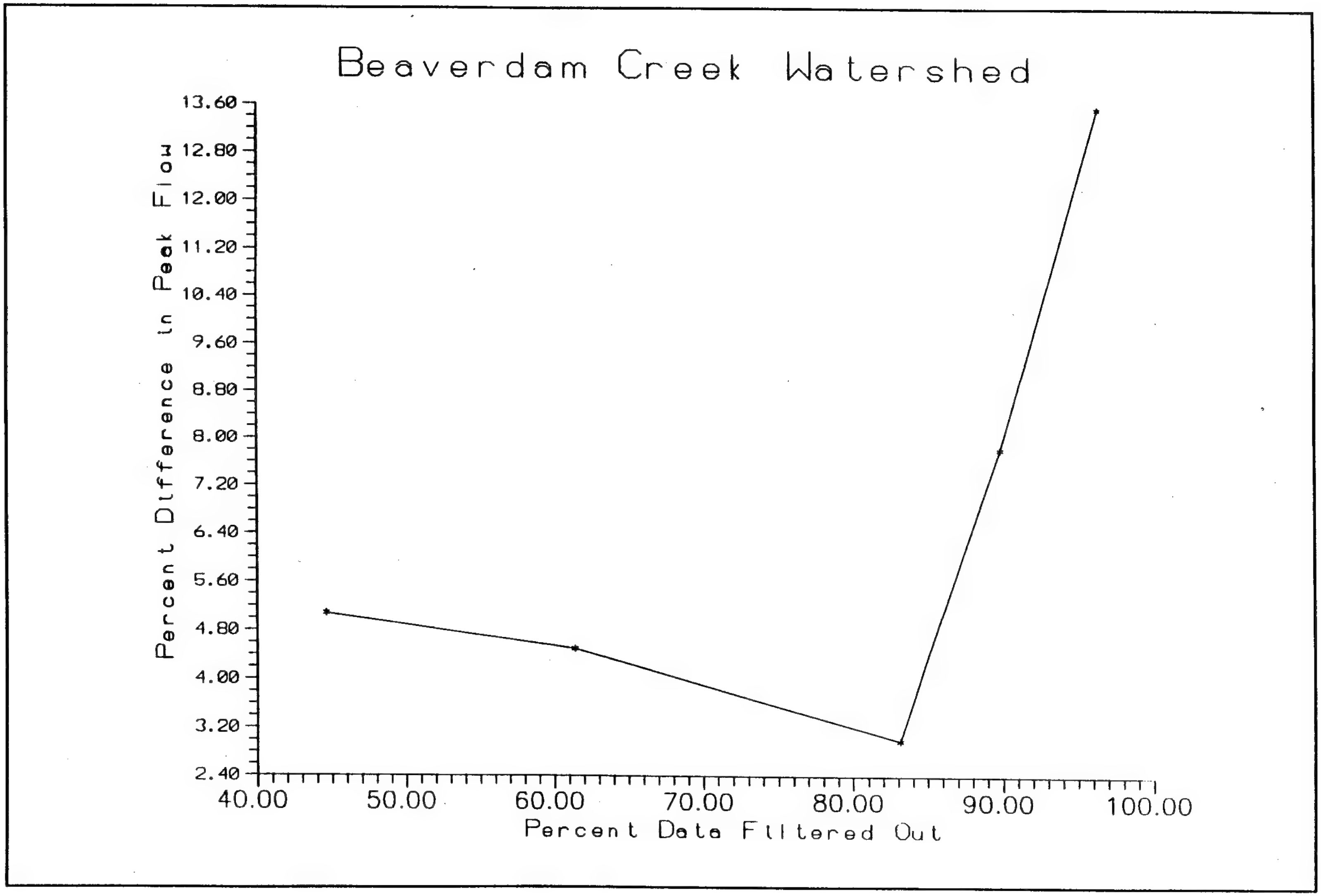



Figure 43. Peak Flow Changes vs. Data Filtered for Beaverdam Creek

differences in the runoff data do not vary significantly until more than about 80% of the data in the DEM grid has been filtered out. This amount of data filtering corresponds to a data filtering threshold of 6 or 7. Logic dictates that the results should not improve as more data are filtered to create a TIN, as seen for Rock Creek and Beaverdam Creek. In fact, the improvements in peak flow in those cases are so slight as to be considered relatively insignificant.

With this information, guidance on the impact of filtering of elevation data on the results of a hydrologic model can be offered. No single geometric parameter accounts for the runoff from a basin, but it is the combination of various parameters that determine the runoff. By simply viewing the geometric data presented in Chapter VIII, it is difficult to determine the impact on model results. However, when the combined effects of those data are accounted for through the model, the impacts become apparent. When more than approximately 80% of the data are filtered from the original DEM grid using the curvature based method, the peak flow can differ significantly from the unfiltered case. The hydrologist should keep this in mind when using filtered elevation data to create a TIN for use in hydrologic modeling.

# CHAPTER X

# APPLICATION OF METHOD

As with many theoretical or research topics, focus can easily remain solely on the results of hypothetical cases or sample data sets and fail to demonstrate that the particular method is successful in practical applications. This was not entirely the case for this study in that the three sample data sets used for analysis of the curvature based filtering method were taken from actual DEM files. Additionally, the basin delineation, geometric computations, and hydrologic model runs were all performed on actual watershed areas within those DEM files. However, even though actual data for the watersheds were used, the rainfall and loss rate parameters were arbitrary and the model results were not compared to observed data. To add validity to the usefulness of this data filtering process and the subsequent use of TIN derived data in hydrologic modeling, an actual project on which this technique was employed is detailed in this chapter.

## Project Background

This demonstration project involved acquisition of DEM data for the Sycamore Creek watershed in central Arizona,

filtering those data with the curvature based technique, generating a TIN from the filtered data, delineating the watershed and subbasins, computing the geometric data from the TIN, setting up an HEC-1 model, and comparing model results with observed rainfall and streamflow data. The Rock Creek basin, which served as one of the sample data sets in the research detailed in this report, is a small portion of the Sycamore Creek watershed. The area of the Rock Creek basin is approximately 65 square kilometers, while the entire Sycamore Creek watershed encompasses approximately 505 square kilometers.

## Elevation Data

The Sycamore Creek watershed extends into portions of two different 1 Degree DEMs. The southern portion is covered by the Mesa West DEM while the northern portion is contained within the Holbrook West DEM. Thus, rectangular portions of each of those DEM files were cropped out, filtered, and joined together to generate one DEM grid file. The unfiltered DEM grid consisted of 720 columns and 900 rows of data, or 648,000 elevation data points, at a 3 arc-second spacing.

## Data Filtering and TIN Generation

The curvature based filtering method detailed earlier in this report was used to select a set of surface specific points from the DEM grid to create a TIN. The first

question to be answered during this process was how to determine the proper filtering threshold. Here, the guidance provided by this research proved useful. Based upon the results of the hydrologic modeling presented in Chapter IX, the desired filtering threshold should remove no more than approximately 80% of the data points from the DEM grid. In fact, a removal of about 80% should result in the optimum data set, i.e., one containing the least number of data points while still providing topographic data that will result in an accurate hydrograph. As shown in Chapter VI, the filtering threshold of 6 resulted in 83% of the DEM data being filtered out for the Rock Creek basin, so a threshold of 6 was used for this project. After filtering the DEM grid elevation data for Sycamore Creek at the 6 threshold, 109,628 data points out of the original 648,000 were retained, which means that 83.1% of the data points were removed. Filter thresholds of 7 and 5 were also tried, which resulted in a removal of 92.9% and 72.0% of the data points respectively. The results from Chapter IX indicated that removal of 93% of the data (threshold 7) for Rock Creek resulted in a deviation of almost 10% in the peak flow of the hydrograph while removal of 71.9% (threshold 5) only affected the hydrograph by 0.44%. Therefore, based upon the results of this research, the filter threshold of 6 was considered adequate and sufficient for this project. Using the 109,628 surface

specific points retained during the filtering process, a TIN was created using the Delauney triangulation criterion.

## Watershed Delineation

The outlet of the watershed was identified on the TIN where Sycamore Creek flows into the Verde River. That point was selected as a terminus, and the watershed delineation routine was used to identify all triangles which contributed flow down to that point. The resultant set of triangles which contributed flow to that outlet comprised the Sycamore Creek watershed. Figure 44 shows the exterior boundary of the TIN with the watershed boundary for Sycamore Creek shown in the interior of the TIN. After the watershed boundary was delineated, the triangles in the remainder of the TIN, which fell outside of the watershed boundary, were discarded from the file.

As previously noted, a watershed is typically subdivided into several smaller basins when creating an HEC-1 input file. The reason for this is that the input variables for a hydrologic model of a larger watershed, such as the loss rate, will typically vary throughout the watershed. Additionally, rainfall will not usually fall uniformly over a large area. Therefore, the area is divided into smaller regions which can reasonably be assigned uniform values for such variables as loss rates and rainfall intensity. This was done with the Sycamore Creek watershed with 7 subbasins identified for the model.

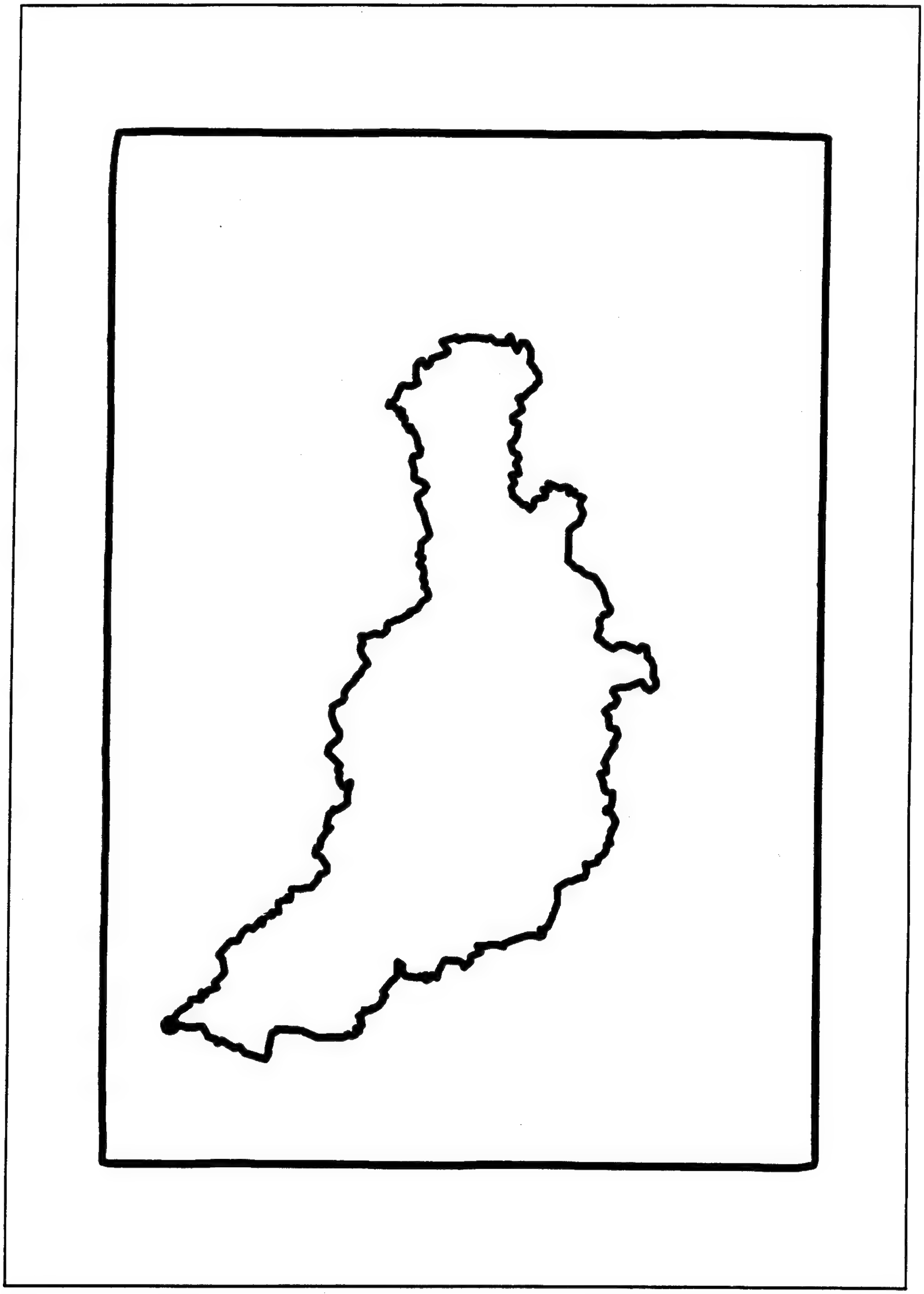


Figure 44. Sycamore Creek Watershed

The outlets of each of those basins were identified as terminus points, and the watershed delineation routine was used again to determine the set of triangles in the TIN that belonged to each subbasin.

## Watershed Geometric Parameters

As with the sample data in the earlier sections of this document, the geometric parameters of the Sycamore Creek watershed were computed based upon the geometry contained in the triangles which comprise the subbasins. The area in square kilometers, average slope, and maximum flow distance in meters were computed for each of the 7 subbasins in the Sycamore Creek watershed. Figure 45 shows the 7 subbasins delineated along with the geometric parameters for each subbasin. Recall from the introduction to this chapter that the area of the Sycamore Creek watershed was listed as approximately 505 square kilometers. This area was determined by delineation of the watershed boundary on 1:24,000 scale USGS quadrangle maps and using a planimeter to measure the area. As a comparison to this manually derived area, a summation of the areas shown for the 7 subbasins in Figure 45 indicates that a total area of 505.90 square kilometers was computed for the Sycamore Creek watershed from this TIN. The difference between these manually computed and TIN derived areas is less than 0.18 percent.

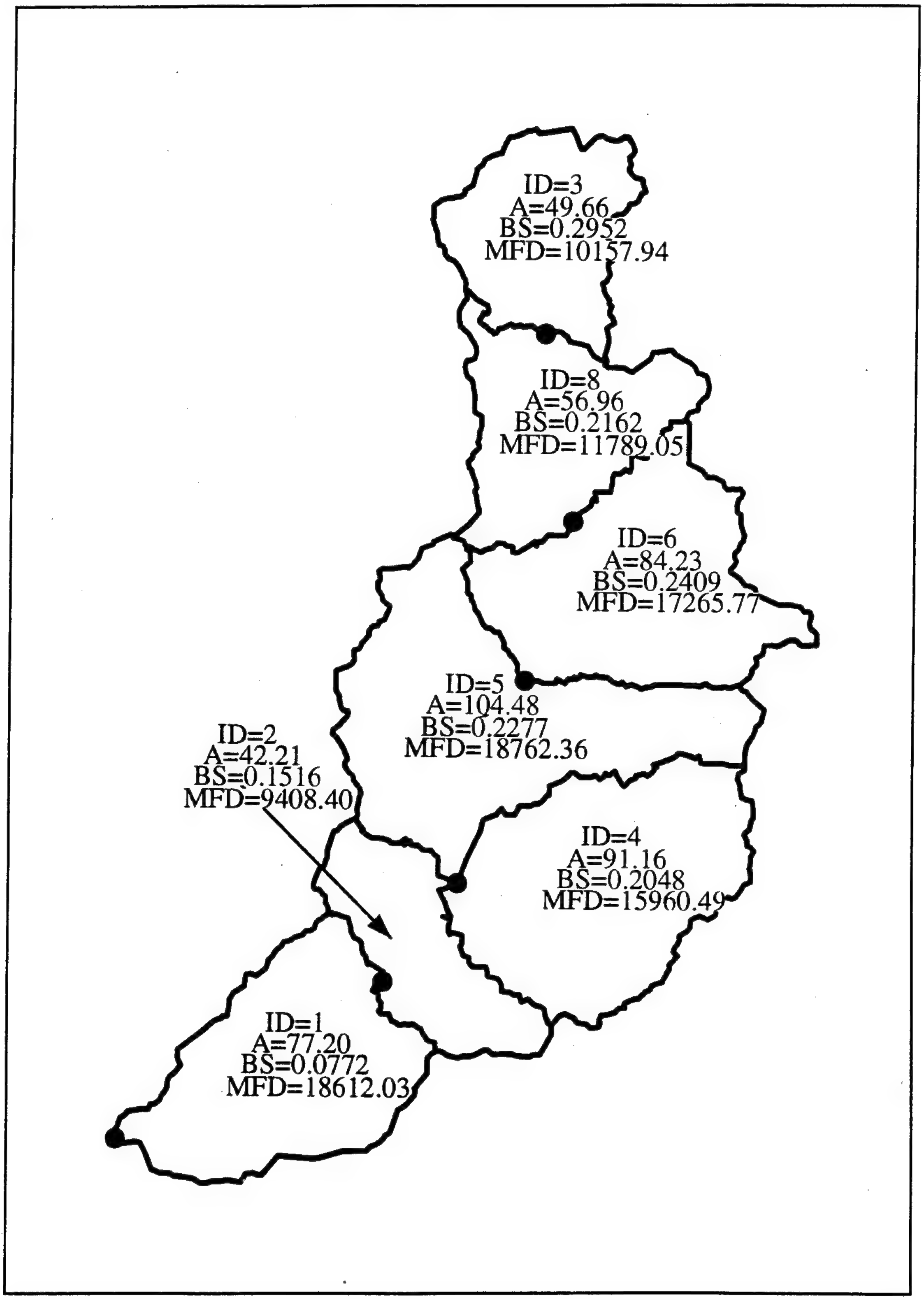



Figure 45. Sycamore Creek Subbasins

## HEC-1 Model Input

As with the sample HEC-1 model runs presented in Chapter IX, an HEC-1 input file was generated using the geometric data shown in Figure 45. The values for basin area and the unit hydrograph parameter TLAG were derived directly from those geometric data. The loss rate parameters were determined based upon a combination of soil type, land use, and land cover over each of the subbasins. Observed rainfall data from several historic events were used.

An additional requirement for this model, which was not included in the test watersheds discussed earlier in this document, was the need to route the flow from the upper basins in the watershed to the outlet. This is another aspect of hydrologic modeling for which a TIN can provide useful information. The modeler can identify flow paths on the TIN which represent the stream network for the watershed being modeled. Based upon the geometry of the TIN along those flowpaths, the length and average slope for each segment of the stream network can easily be computed. Those values can then be used, along with channel cross sectional data and roughness characteristics, as part of the input required for channel routing in the model. The effect of the curvature based filtering method on the lengths and slopes of streams on a TIN were not analyzed in this research, but future research in this area would be useful.

## HEC-1 Model Results

The above mentioned input parameters were used to generate an input file for the HEC-1 model, and the results of the model were compared to observed streamflow data for the watershed. Initial results from the model showed that the computed runoff was higher than the observed flow. Adjustments to the loss rate parameters were made and the model was run again with better results, and the results are shown in Figure 46. This Figure includes the observed and the computed hydrographs for the Sycamore Creek watershed for one of the rainfall events used to verify the model. As can be seen from those hydrographs, the model reproduced the observed data extremely well. Thus, the TIN based terrain model can effectively be used to generate input variables for hydrologic modeling, and the curvature based method for selecting the TIN data points from a DEM is a viable approach.

One additional comparison was attempted for this project in support of the research presented in this thesis. That comparison was to have been between the hydrograph results obtained using data filtered at a threshold of 6, as presented above, and results obtained using the unfiltered DEM grid. This attempted comparison provides another reason why the elevation filtering process is important in TIN modeling. That reason is that the TIN data file generated using the unfiltered DEM grid, which

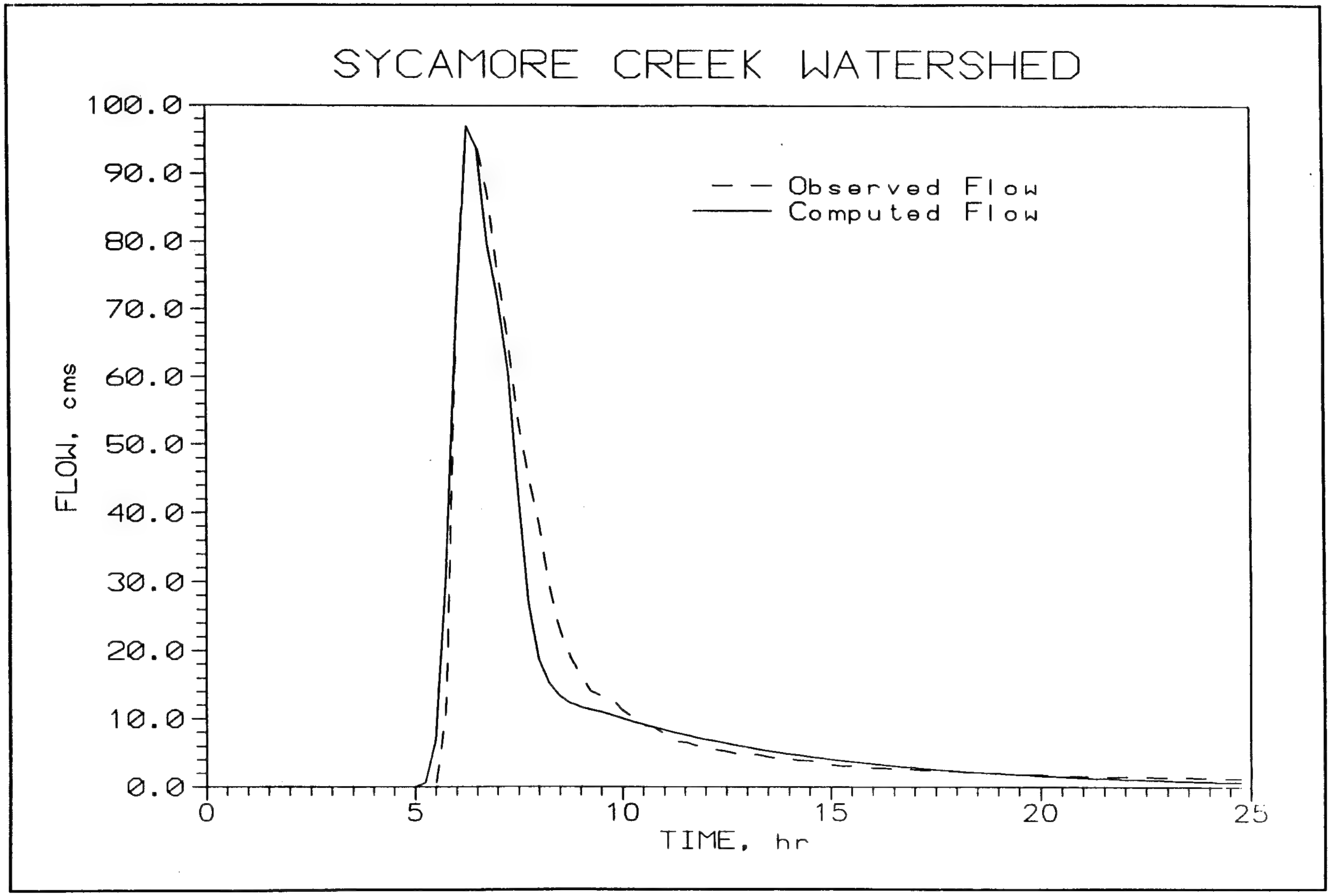



Figure 46. Sycamore Creek Model Results vs. Observed Flow

consisted of 648,000 points, was so large that the triangulation could not be successfully completed. The triangulation was attempted on two different engineering workstations, and there was insufficient memory on each of them to complete the process. The initial data file, prior to the attempted triangulation, contained nearly 20,000,000 bytes of data just for the elevation points. The size of a TIN file typically more than doubles after it has been triangulated, due to the additional information in the file regarding which points make up each triangle. Therefore, if the 648,000 point DEM grid file for Sycamore Creek had been successfully triangulated, it would have contained roughly 50,000,000 bytes of data. A file that size makes the use of the unfiltered DEM grid virtually impossible, as well as undesirable, for larger areas. If the data filtering method presented in this report were not used, then modeling the Sycamore Creek watershed using a TIN would not have been feasible.

# CHAPTER XI

# CONCLUSIONS AND RECOMMENDATIONS

With the guidance provided from this research on appropriate levels of data filtering, the curvature based technique for selecting surface specific points from DEM data to create a TIN model provides an effective and efficient method for using widely available elevation data to generate input for a lumped parameter hydrologic model. Using this method allows the hydrologist to quickly preprocess topographic data to determine many of the geometric parameters required. Determination of those parameters can be extremely time consuming if performed manually. However, many of the required input variables can quickly and accurately be determined from a TIN based model, thus allowing the hydrologist to focus on the hydrologic analysis portion of the project.

## Conclusions

Specific guidance resulting from this research applies to the use of this method in hydrologic modeling. The results provided in Chapter IX indicate that the model results could be significantly affected if more than approximately 80% of the data are filtered out of the

original DEM. With less than 80% removed, the model results did not change significantly from those obtained using the entire original DEM data.

Application of the curvature based filtering technique for the computation of hydrologic model input data using TINs was demonstrated in Chapter X of this report for the Sycamore Creek watershed. A review of the time and effort required to complete a TIN model for the Sycamore Creek watershed clearly demonstrates the efficiency of this method for preparing input data for hydrologic modeling. The process of obtaining elevation data, filtering those data, generating a TIN, delineating the watershed and subbasins, and computing geometric data for the subbasins was done in a matter of a few hours from start to finish. The elevation data used are available via the internet from the USGS and were obtained in approximately 10 minutes. The same elevation product can similarly be obtained for any region in the United States. Once the DEM data were obtained, the initial processing and filtering were completed in about 30 minutes. This included filtering the data at three different threshold levels, determining the amount of data removed at each threshold, and selecting the appropriate filtering level. After completion of the filtering, the TIN was generated in less than 5 minutes. The only portion of the process that was somewhat time consuming was the delineation of the watershed and subbasin

boundaries. There are typically some areas in any TIN which require minor editing and additional processing before the watershed boundary can be accurately determined. For the Sycamore Creek watershed, this required about 4 hours of time. After the boundaries were determined, the computation of geometric parameters required only a few minutes. Thus, starting with no data other than a knowledge of the location of the watershed, the DEM data were obtained; the DEM data were filtered; the TIN was created; the watershed and subbasins were delineated; and the geometric parameters for the model were computed all in less than 6 hours from start to finish. Obviously, this process would take many times longer if it were done manually. This example clearly demonstrates that by using this method, a very small portion of the entire time spent on the modeling effort was spent determining the watershed geometry, with the majority of the effort spent on other aspects of the hydrologic analysis.

Although the curvature based filtering method is effective and simple to use, the user should have some knowledge of the process to select the appropriate filtering threshold for a given set of data and for a given purpose. The creation of a TIN using this filtering method could be used for several different reasons. Although this research focussed on the use of TINs for hydrologic modeling, a TIN may also be used simply for visual display

of topography, for determining a watershed boundary, for estimating only one specific geometric parameter of the topography such as slope or aspect, or for several other purposes. For each of these uses, the required quality of the TIN to produce acceptable results will probably differ. This research included analysis of the effects of the curvature based method on the statistical results, geometric parameters, and hydrologic modeling results. Thus, the results presented in this report should provide useful guidance for any of the above mentioned applications of TIN modeling.

## Recommendations

Although the filtering method as presented is effective, there are enhancements to the technique that should be considered and additional research that should be pursued. One enhancement would be to allow the user to specify a particular tolerance for the error between the TIN and the original DEM data. The selection of a particular filtering threshold does not guarantee that the error will be within a given tolerance. Thus, if there is a specific requirement for a tolerance to be met, it would be useful to allow the user to specify that tolerance. The same is true if the user has a particular need to create a TIN with a specific number of data points. One of the advantages of the method is that the specification of an error tolerance or a number of points is not required, but

there may be instances when a tolerance or number of points is required and the ability to specify those criteria would be useful.

In addition to potential enhancements to this particular method, additional research should be pursued. One area of research would be to conduct a direct comparison of the curvature based filtering method with other existing methods described by Lee (1991a and 1991b). Another potential area of research would be to study the effects of data filtering on the determination of stream lengths and slopes in a watershed. As mentioned in Chapter III, a flowpath on a TIN can be used to simulate a stream, and the length and slope of that stream can be computed from the TIN. However, that length and slope will certainly differ when a different number and configuration of triangles are used. Therefore, analysis of how those computations may be affected by different filtering thresholds would be a useful research area. By pursuing these additional research areas, additional insight into this simple yet powerful method of selecting points for TIN generation will be gained.

REFERENCES

Chen, Zi-Tan, and Guevera, J.A. (1987). "Systematic Selection of Very Important Points (VIP) from Digital Terrain Model for Constructing Triangular Irregular Networks," In Nicholas R. Chrisman (Ed.), Auto Carto 8, Proceedings of the Eighth International Symposium on Computer-Assisted Cartography, 50-56.

Chow, V.T., Maidment, D.R., and Mays, L.W. (1988). Applied Hydrology, McGraw-Hill Publishing Company, New York, NY.

DeFloriani, L., Falcidieno, B., Pienovi, C., and Nagy, G. (1985). "Efficient Selection, Storage, and Retrieval of Irregularly Distributed Elevation Data," Computers and Geosciences, 11(16), 667-673.

DeFloriani, L., Falcidieno, B., Pienovi, C., Allen, D., and Nagy, G. (1986). "A Visibility Based Model for Terrain Features," Proceedings of the 2nd International Symposiaum on Spatial Data Handling, Ohio: International Geographical Union Commission on Geographical Data Sensing and Processing, 235-250.

Department of the Army (1993), Flood-Runoff Analysis, Engineering Manual 1110-2-9021, Corps of Engineers, Washington, DC.

Jones, Norman L. (1990). "Solid modelling of earth masses for applications in geotechnical engineering," PhD dissertation, University of Texas at Austin, Austin, TX.

Jones, Norman L., and Nelson, J. (1992a). "Drainage Analysis Using Triangulated Irregular Networks," In Proceedings of the 8th Conference on Computing in Civil Engineering and Geographic Information Systems Symposium, 719-726, American Society of Civil Engineers.

Jones, Norman L., Wright, Stephen G., and Maidment, David R. (1990). "Watershed Delineation with Triangle-Based Terrain Models," Journal of Hydraulic Engineering, 116(10), 1232-1251, American Society of Civil Engineers.

Lawson, C.L. (1977). "Software for $C^1$ surface interpolation," Mathematical Software III, J.R. Rice, ed., Academic Press, New York, NY, 164-194.

Lee, D.T., and Schacter, B.J. (1980). "Two algorithms for constructing a Delauney triangulation," International Journal of Computing and Information Science, 9(3), 219-242.

Lee, J.(1989). "A Drop Heuristic Method for Extracting Irregular Networks for Digital Elevation Models," Proceedings of GIS/LIS '89, Orlando, FL, November 1989.

Lee, J. (1991a). "Assessing the Existing Methods for TIN Extraction," In Baltimore 1991: Technical Papers, 1991 ACSM-ASPRS Annual Convention, Volume 2, Cartography and GIS/LIS.

Lee, J. (1991b). "Comparison of Existing Methods for Building Triangular Irregular Network Models of Terrain from Grid Digital Elevation Models," International Journal of Geographical Information Systems, 5(3), 267-285.

Miller, Irwin, Freund, John E., and Johnson, Richard A. (1990). Probability and Statistics for Engineers, Prentice Hall, Englewood Cliffs, NJ.

Moore, I.D., Grayson, R.B., and Ladson, A.R. (1993). "Digital Terrain Modelling: A Review of Hydrological, Geomorphological, and Biological Applications," In Terrain Analysis and Distributed Modelling in Hydrology, 7-34, John Wiley & Sons, New York, NY.

Nelson, E. James, Jones, Norman L., and Miller A.W. (1994). "Algorithm for Precise Drainage-Basin Delineation," Journal of Hydraulic Engineering. 120(3), 298-312, American Society of Civil Engineers.

Peucker, Thomas K., Fowler, Robert J., Little, James J, and Mark, David M. (1976). Digital Representation of Three-Dimensional Surfaces By Triangulated Irregular Networks (TIN). Technical Report No. 10, Office of Naval Research, Geography Programs, Arlington, VA.

Southard, David A. (1990). "Piecewise Planar Surface Models from Sampled Data," In Scientific Visualization of Physical Phenomena, 667-680, Springer-Verlag, New York, NY.

U.S. Army Corps of Engineers. (1990). HEC-1 Flood Hydrograph Package User's Manual, Hydrologic Engineering Center, Davis, CA.

U.S. Department of Agriculture. (1985). National Engineering Handbook, Section 4, Hydrology (NEH-4), Soil Conservation Service.

U.S. Geological Survey (USGS). (1987). Digital Elevation Models - Data Users Guide. U.S. Department of Interior, Reston, VA.

Watson, D.F. (1981). "Computing the n-dimensional Delauney Tesselation with application to Voronoi polytopes," The Computing Journal, 8(2), 167-172.

Watson, D.F., and Philip, G.M. (1984). "Systematic triangulations," Computer Vision, Graphics, and Image Processing, 26, 217-223.

Wolock, David M., and Price, Curtis V. (1994). "Effects of Digital Elevation Model Map Scale and Data Resolution on a Topography-Based Watershed Model," Water Resources Research, 30(11), 3041-3052.

# APPENDIX A

# PLOTS OF FILTERED ELEVATION DATA

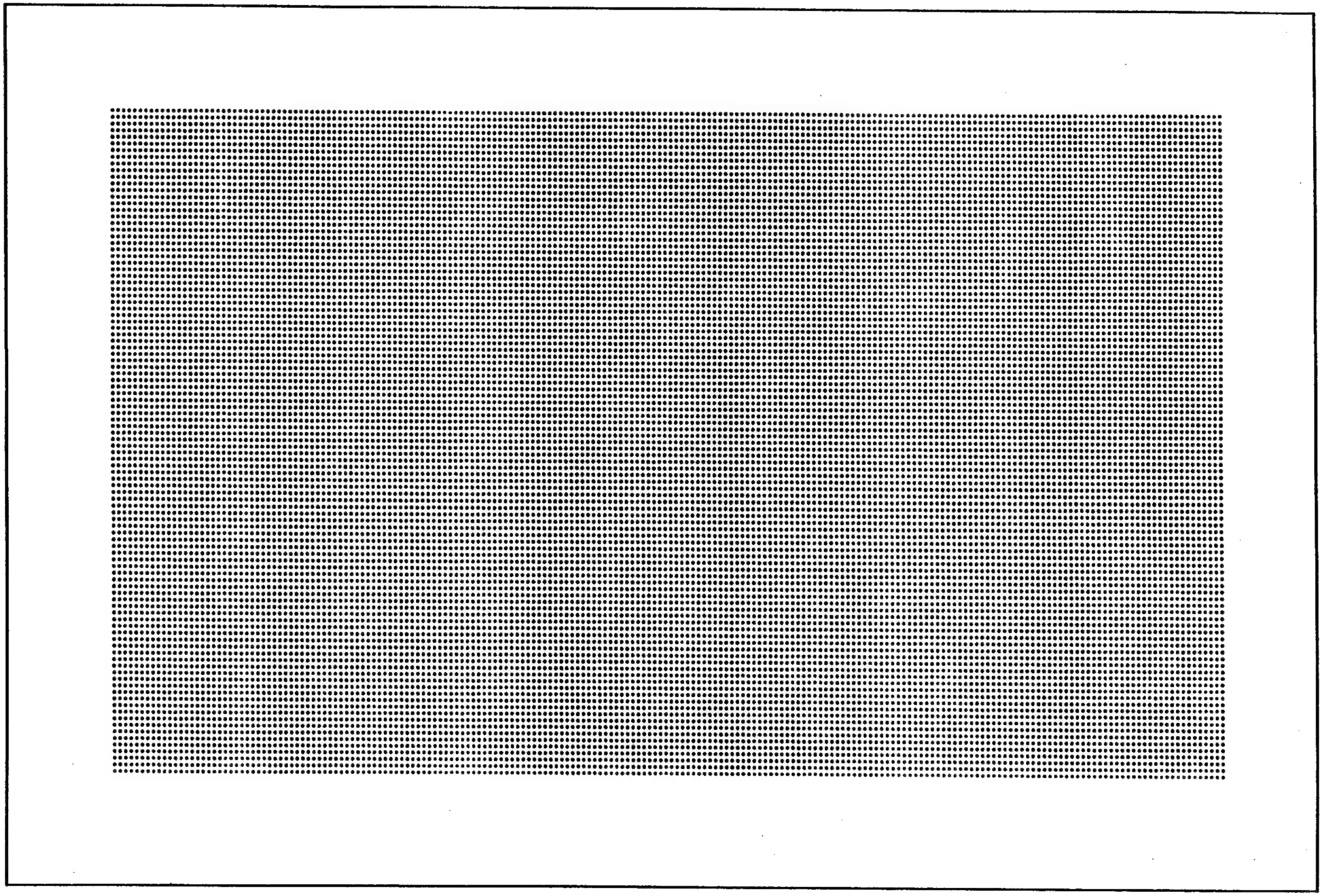


Figure A1.  Rock Creek DEM Data - No Filtering

Figure A2.  Rock Creek DEM Data - Filter Threshold 0

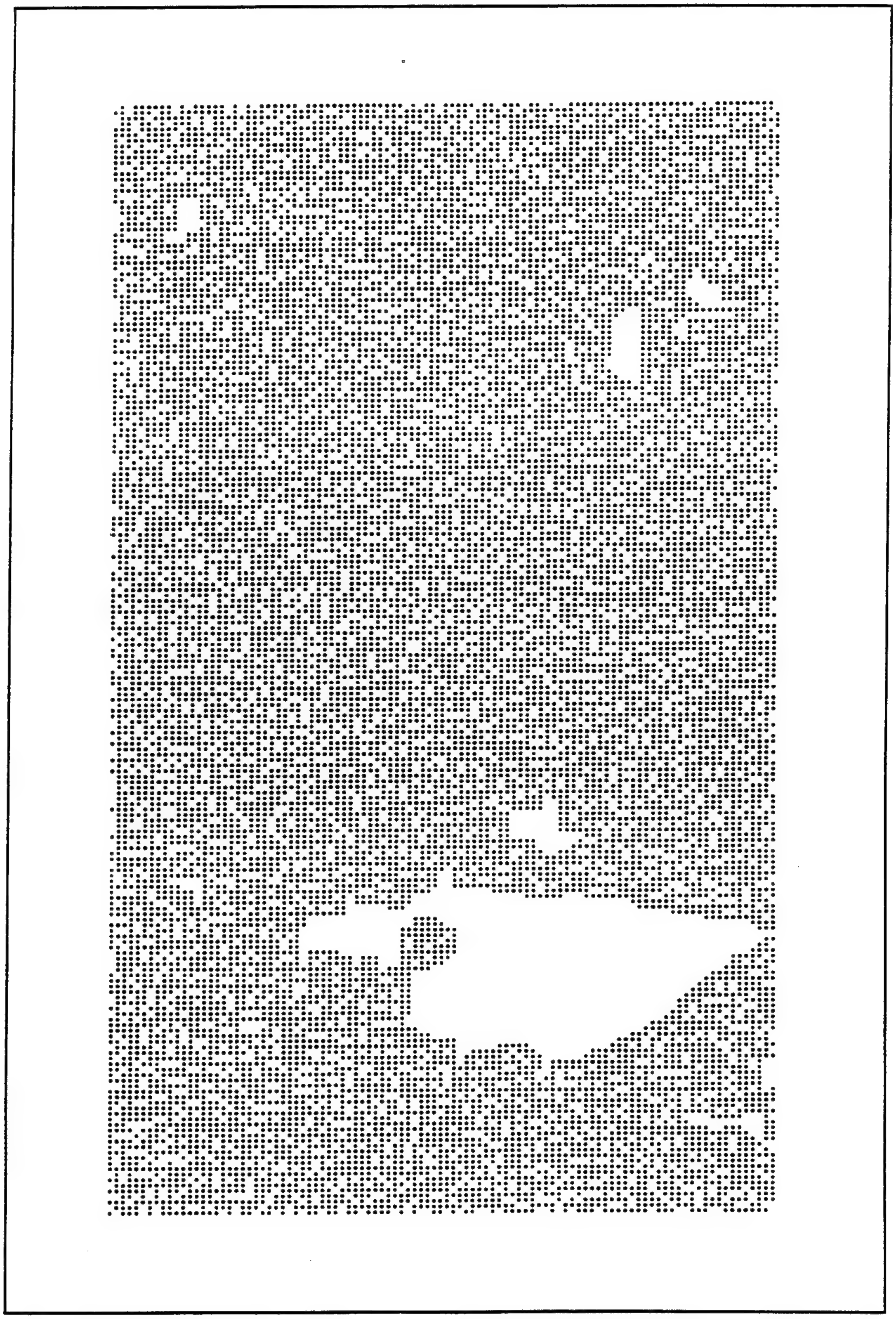


Figure A3. Rock Creek DEM Data - Filter Threshold 1

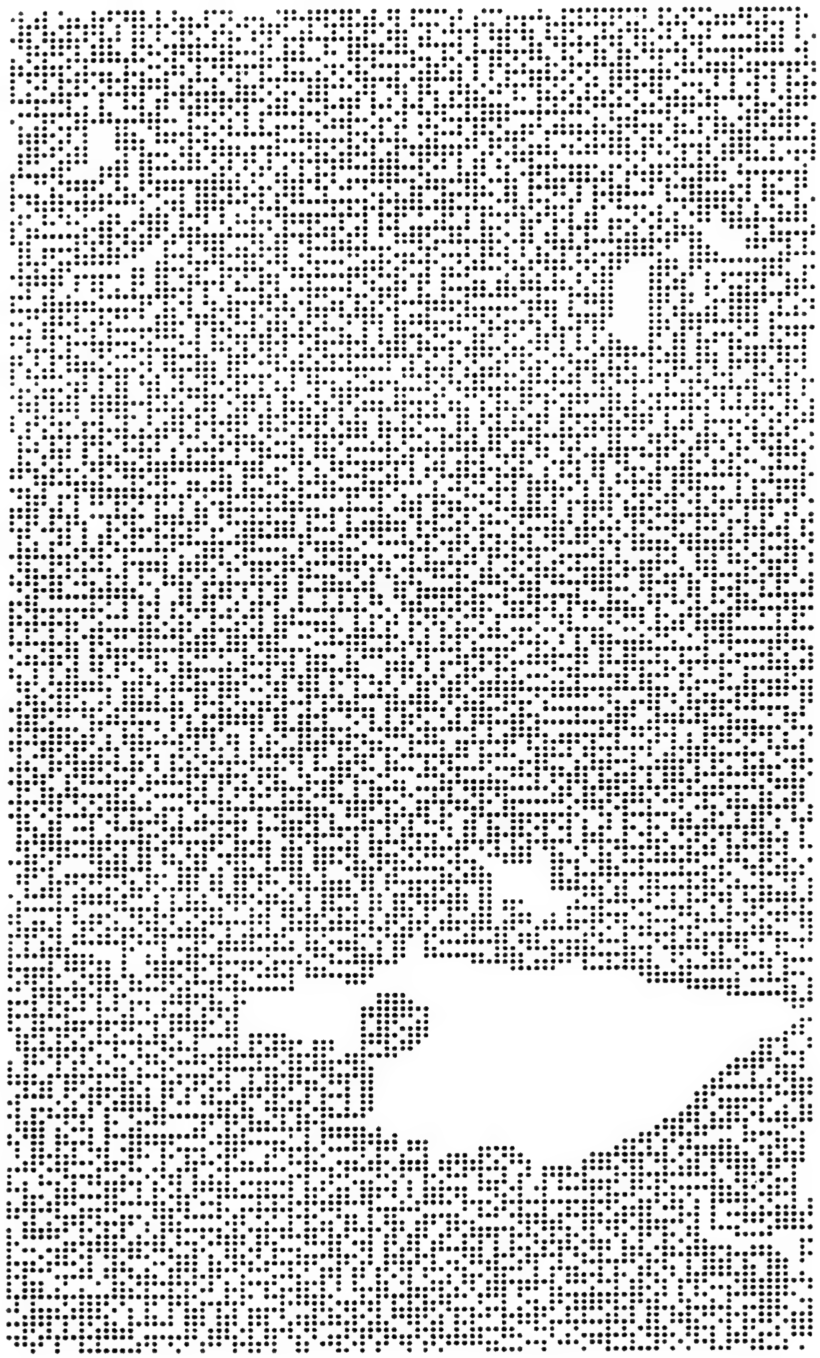


Figure A4. Rock Creek DEM Data - Filter Threshold 2

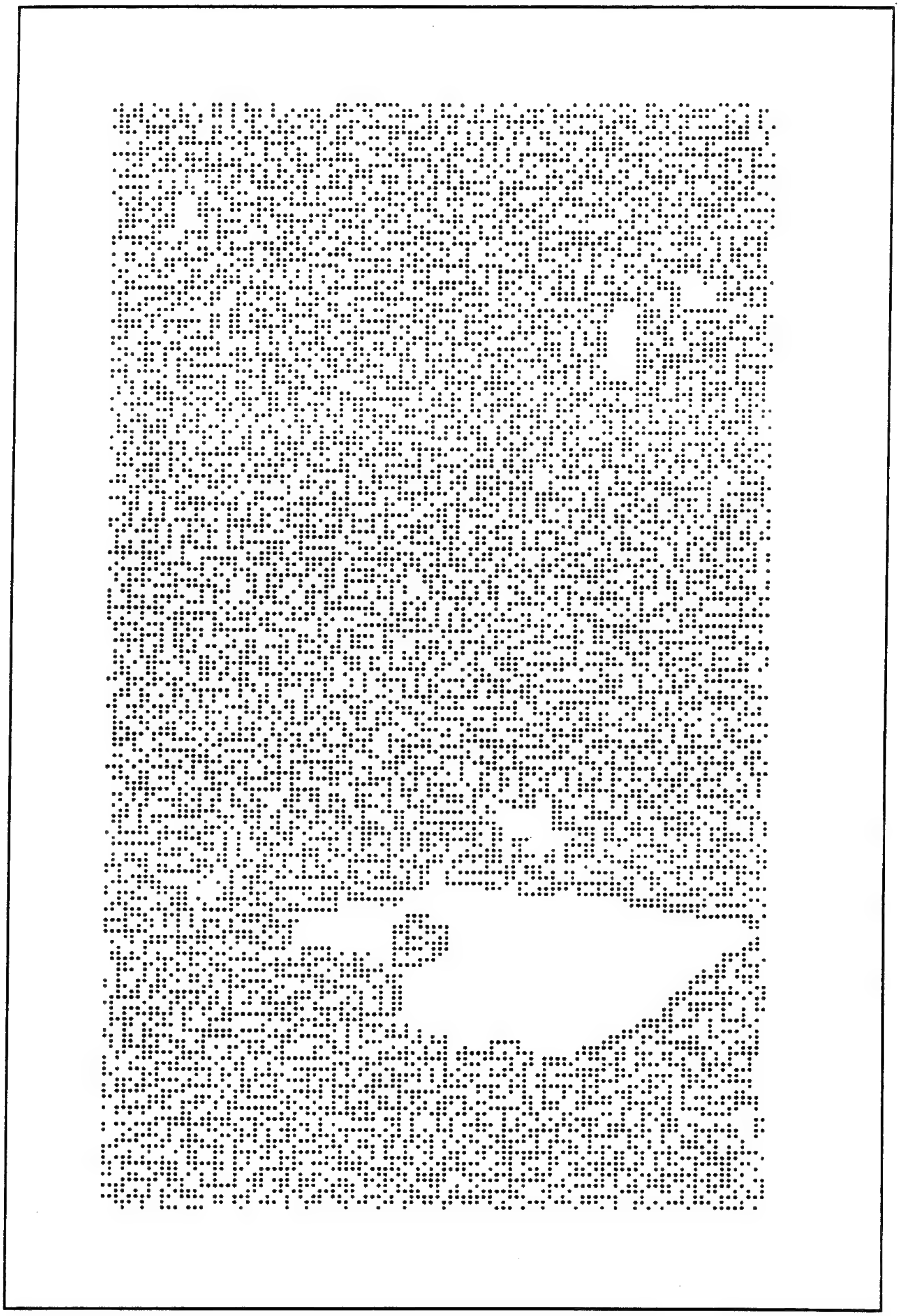


Figure A5.  Rock Creek DEM Data - Filter Threshold 3

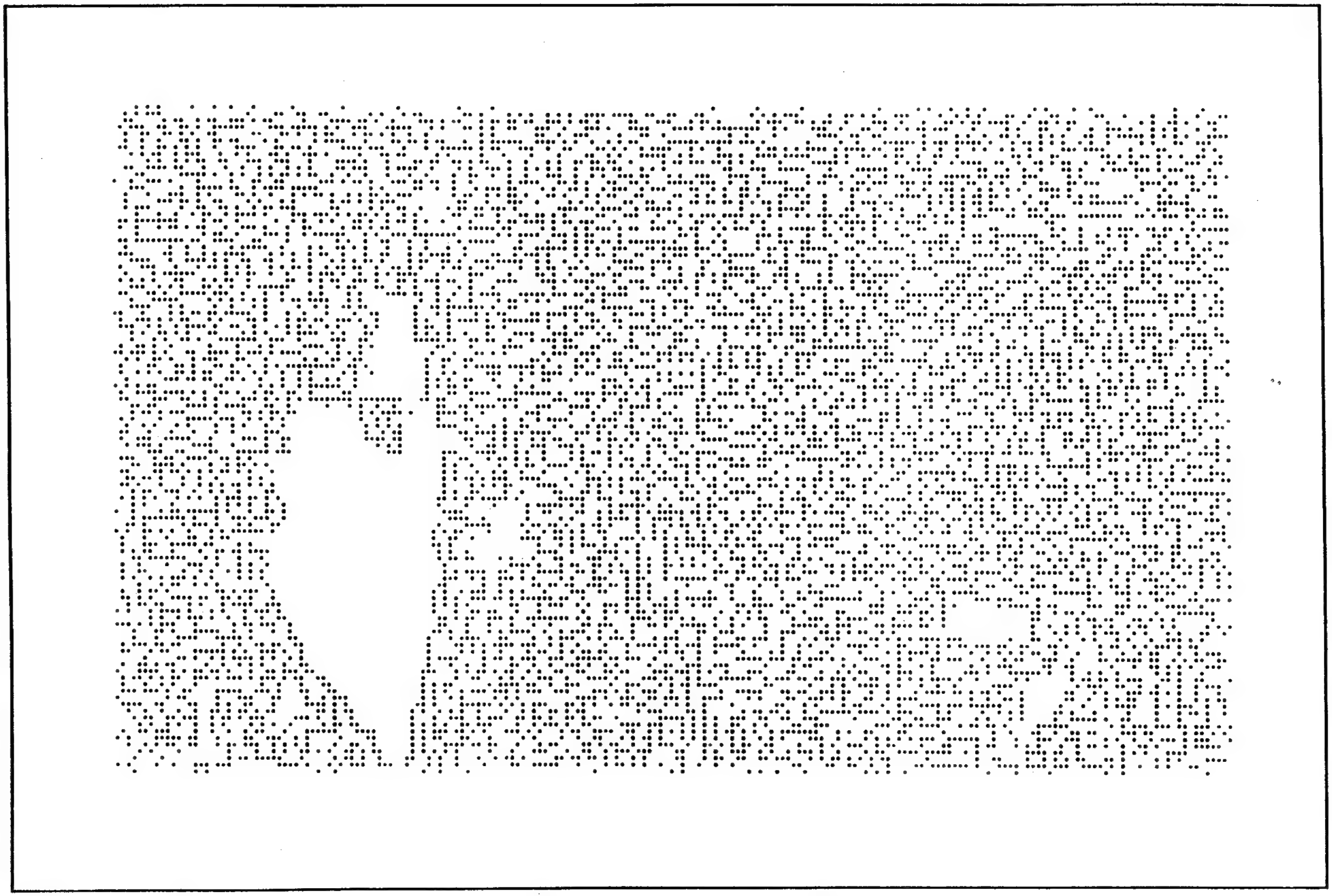


Figure A6. Rock Creek DEM Data - Filter Threshold 4

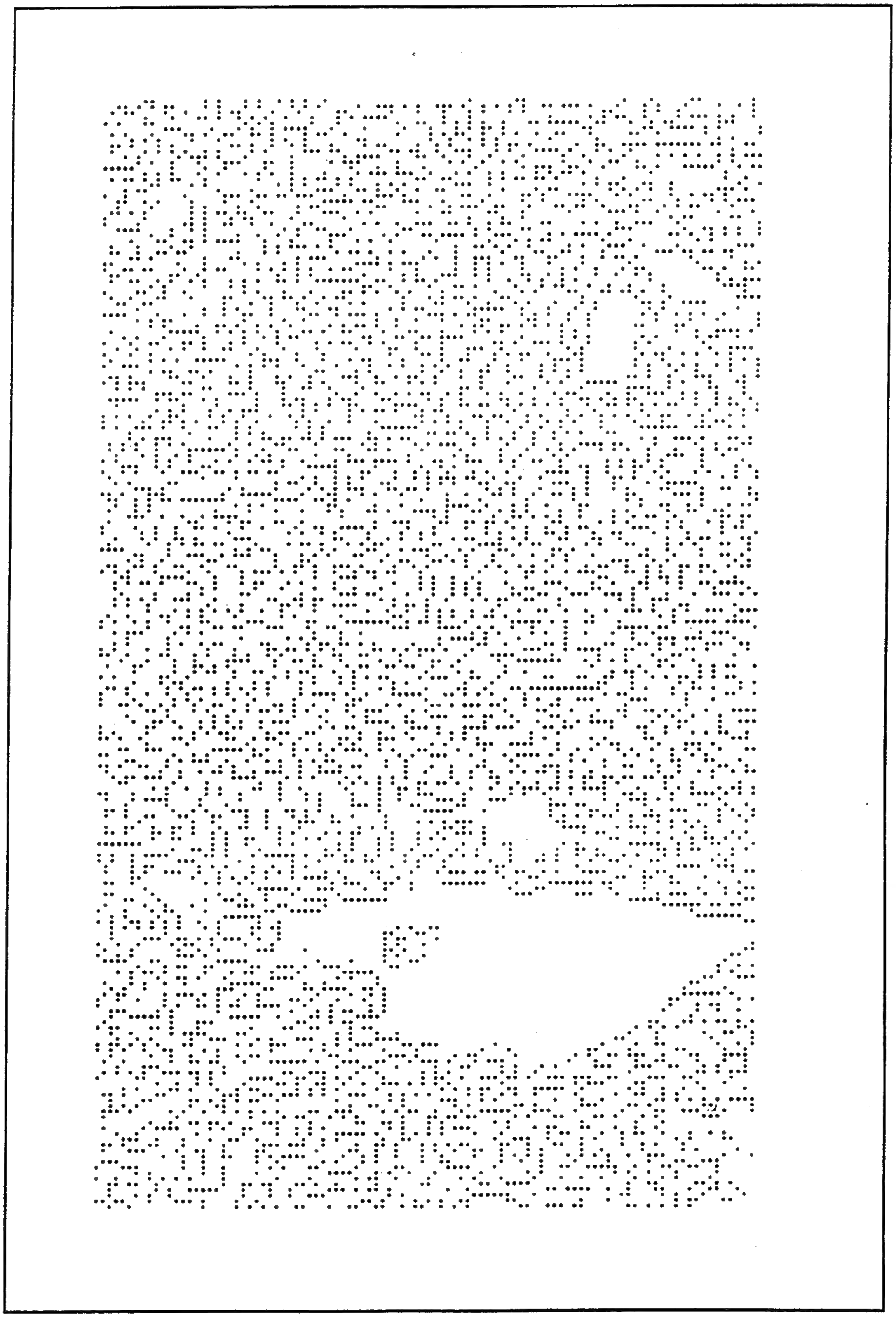


Figure A7. Rock Creek DEM Data - Filter Threshold 5

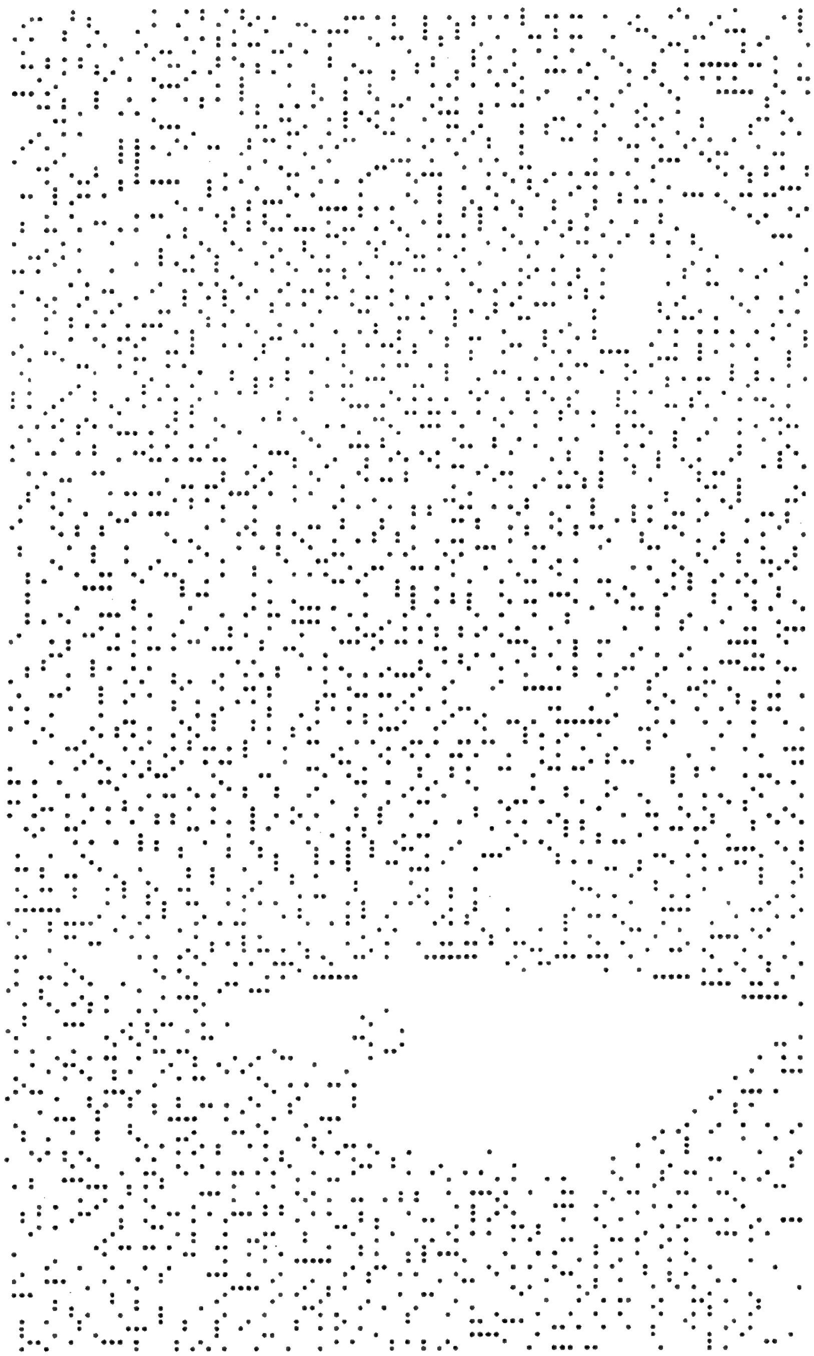


Figure A8. Rock Creek DEM Data - Filter Threshold 6

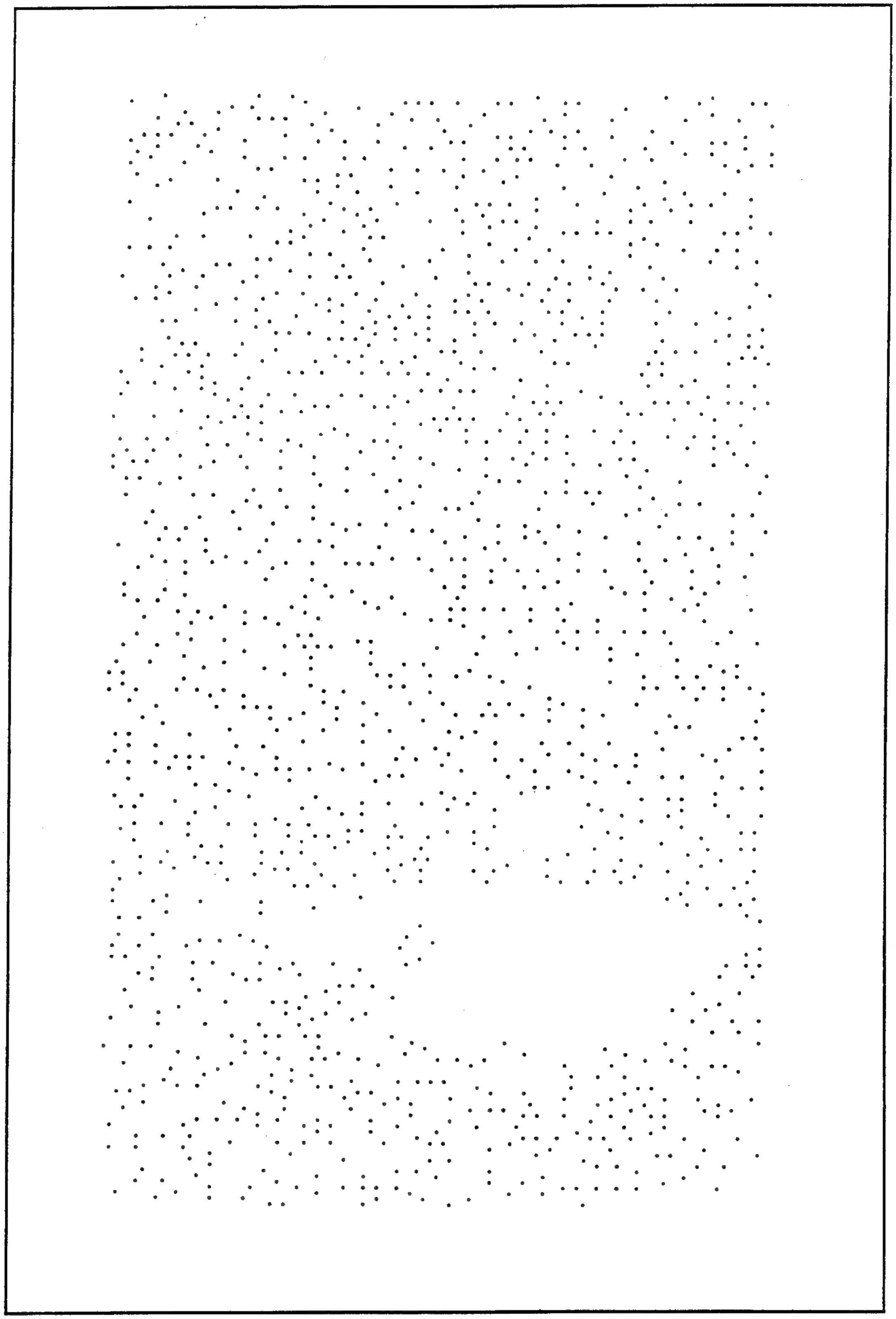


Figure A9.  Rock Creek DEM Data - Filter Threshold 7

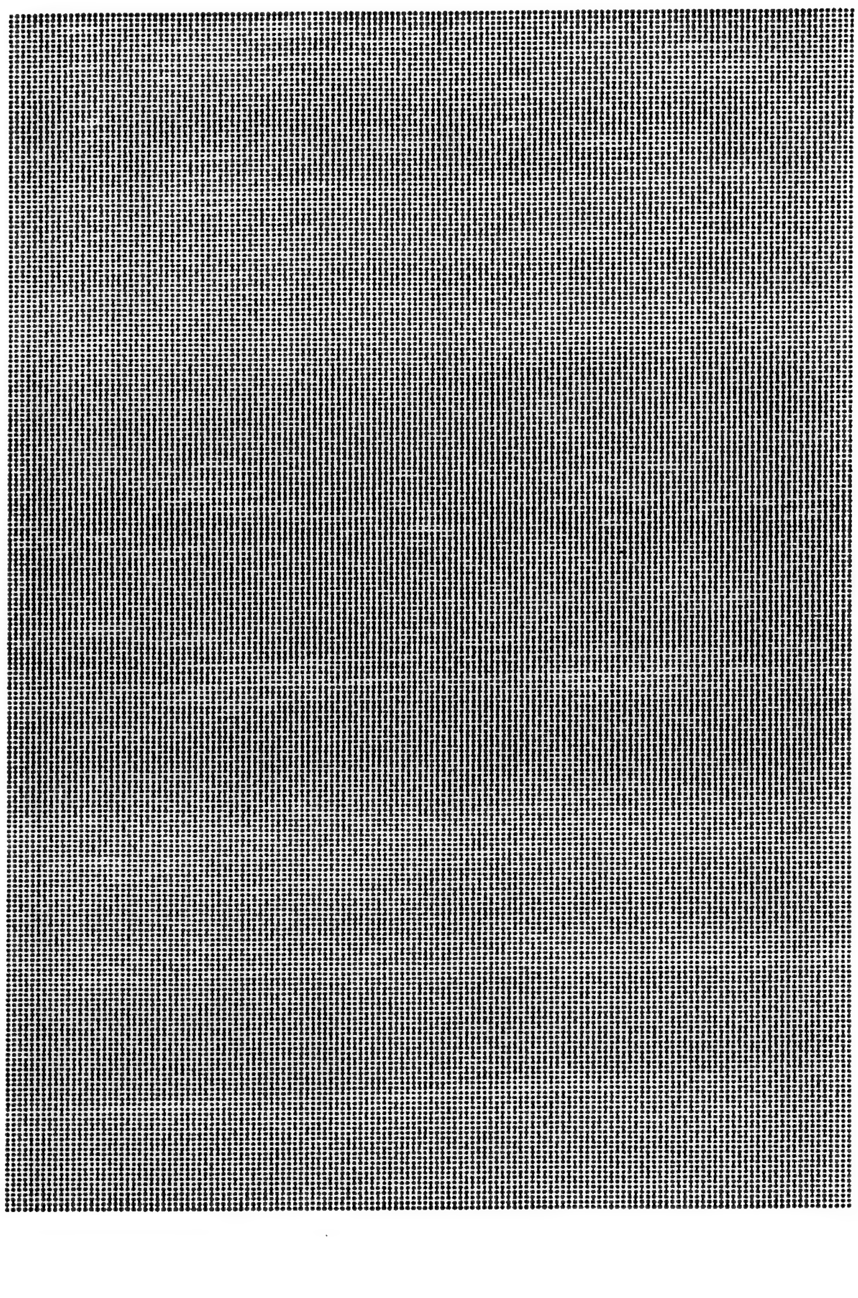


Figure A10. Long Creek DEM Data - No Filtering

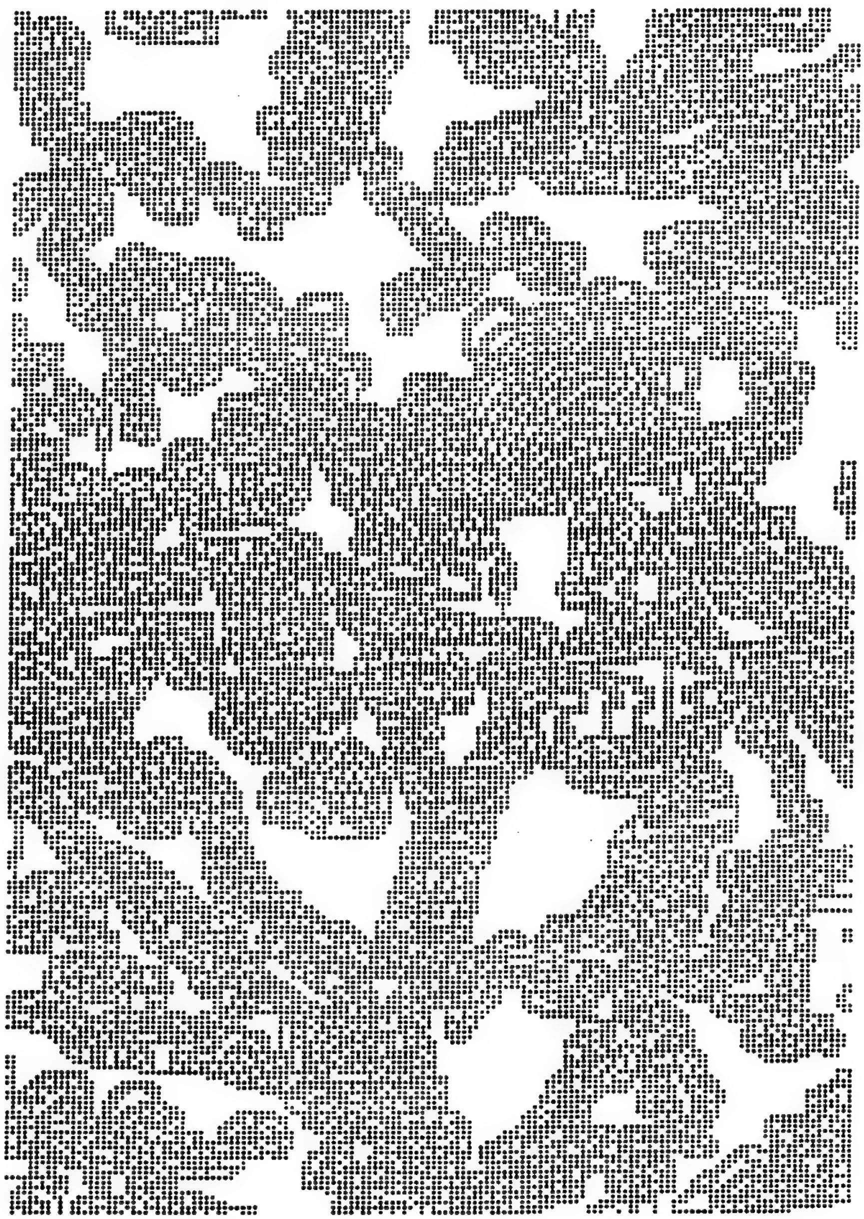


Figure A11. Long Creek DEM Data - Filter Threshold 0

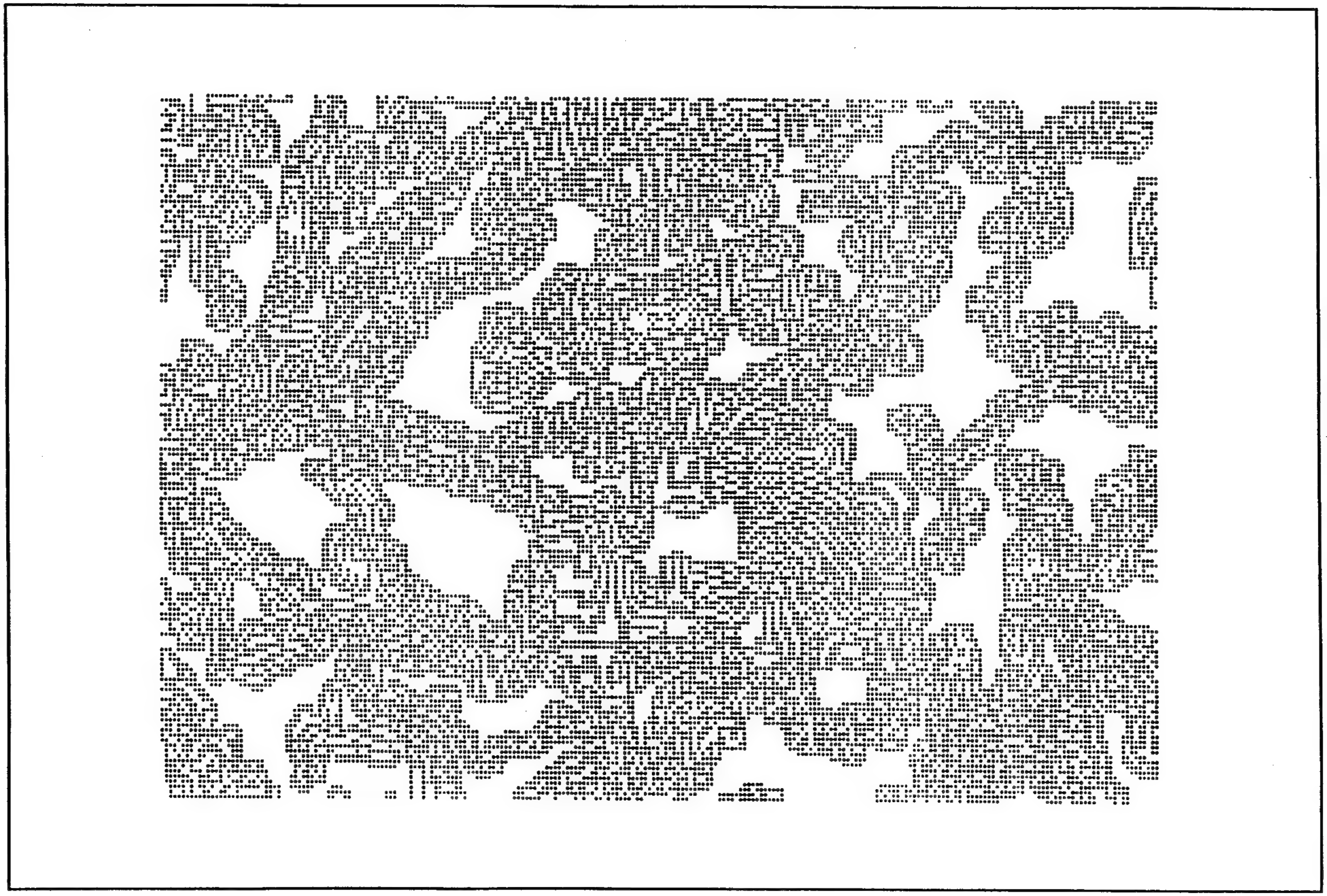


Figure A12.  Long Creek DEM Data - Filter Threshold 1

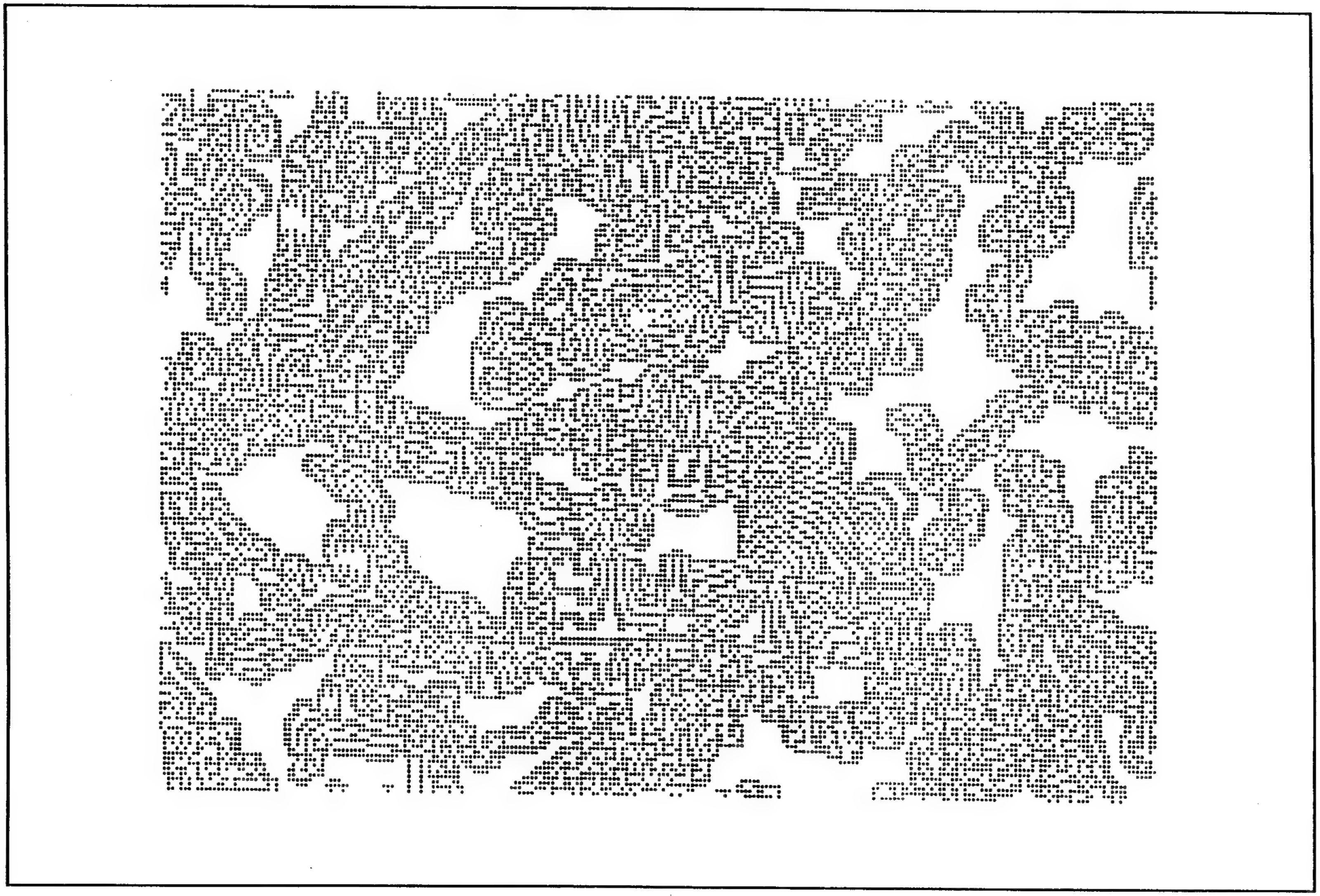


Figure A13. Long Creek DEM Data - Filter Threshold 2

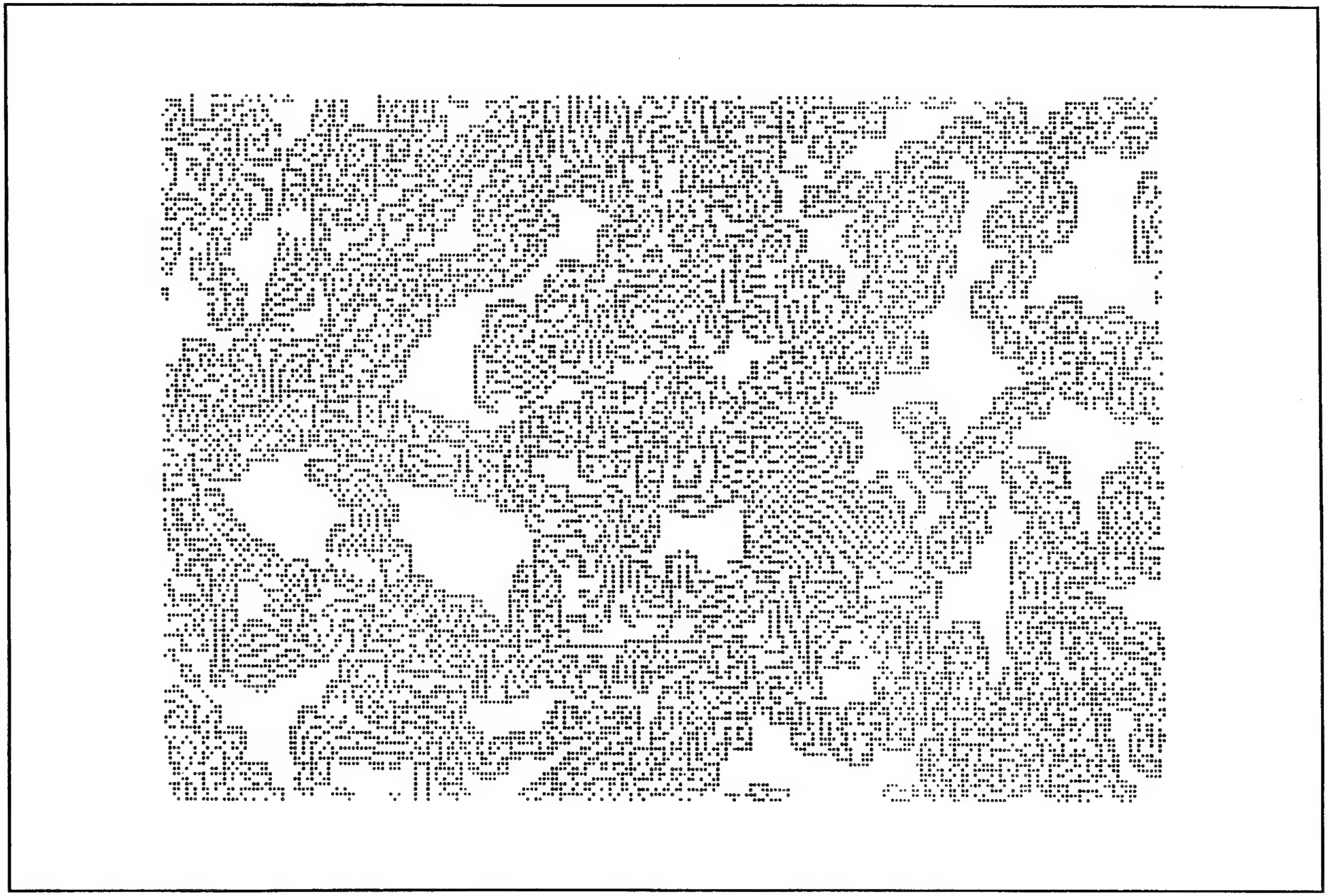


Figure A14. Long Creek DEM Data - Filter Threshold 3

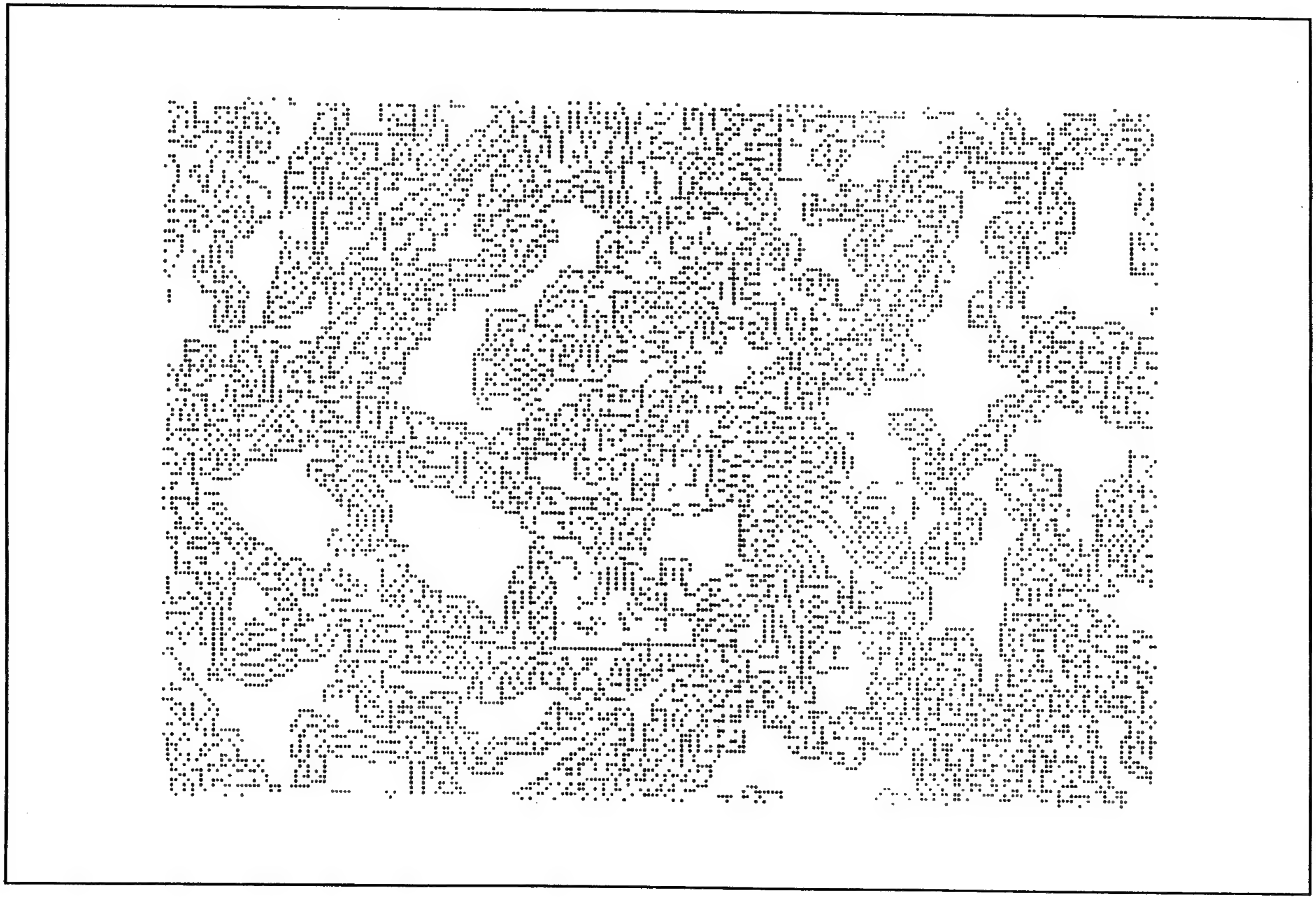


Figure A15. Long Creek DEM Data - Filter Threshold 4

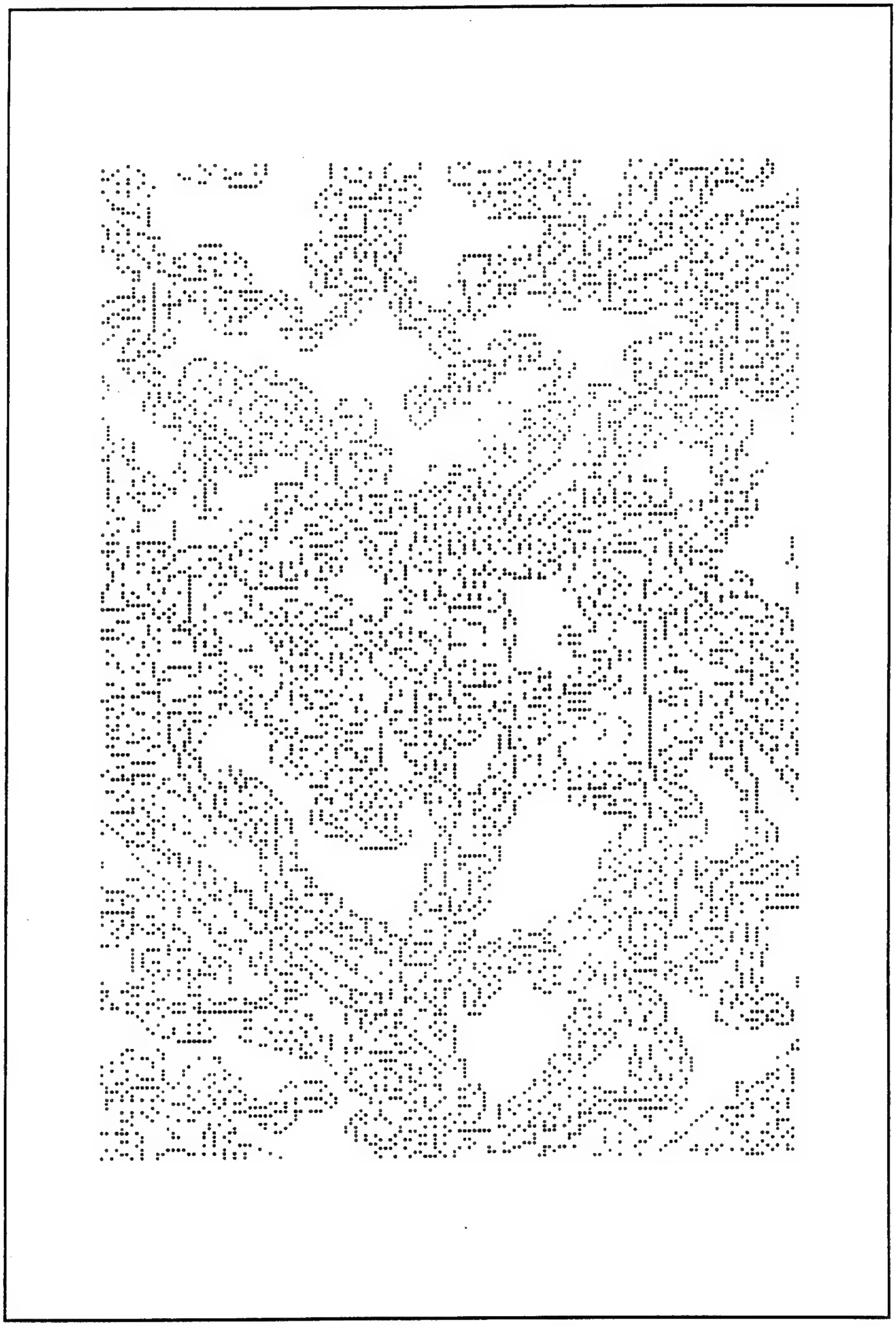


Figure A16. Long Creek DEM Data - Filter Threshold 5

Figure A17. Long Creek DEM Data - Filter Threshold 6

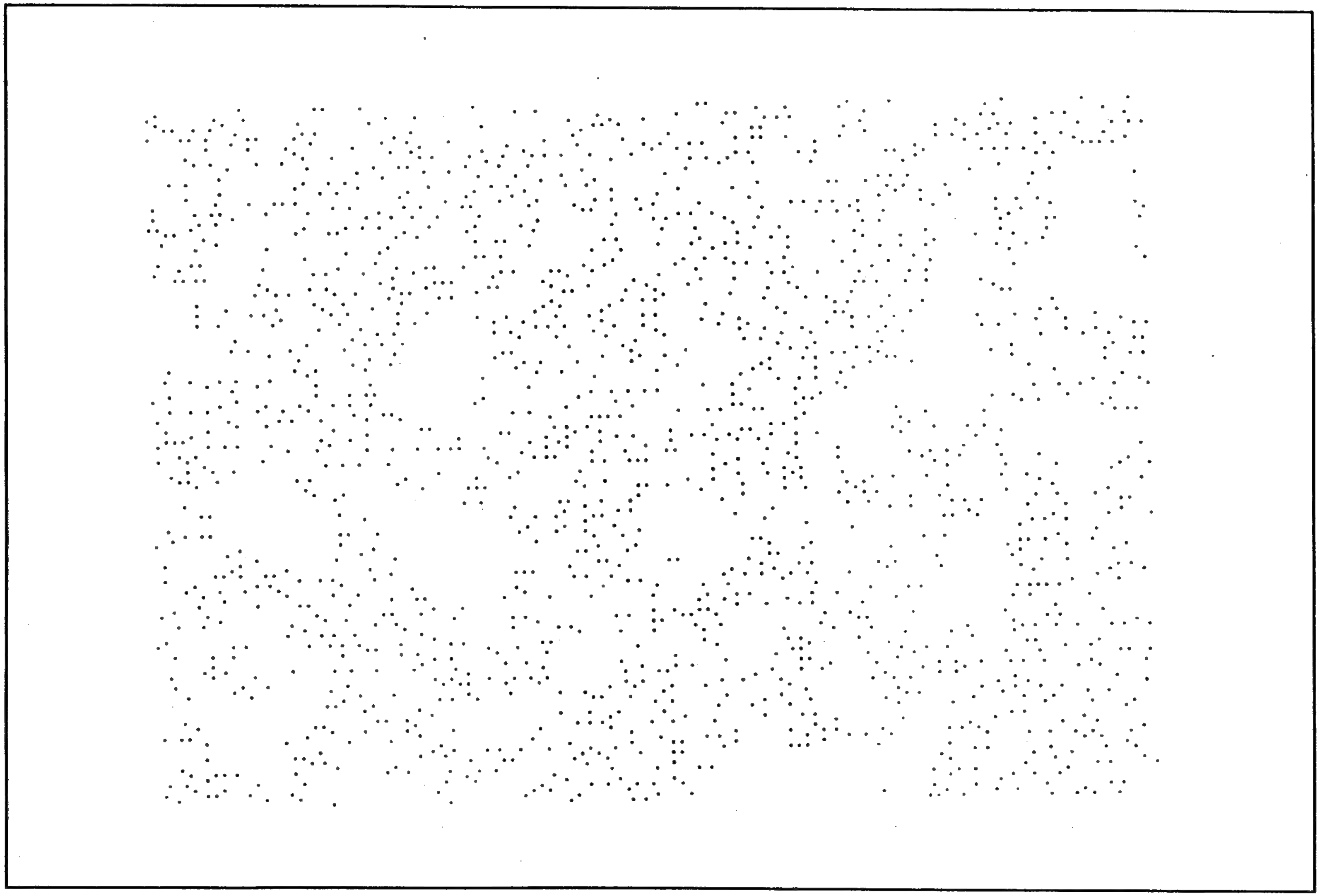


Figure A18.  Long Creek DEM Data - Filter Threshold 7

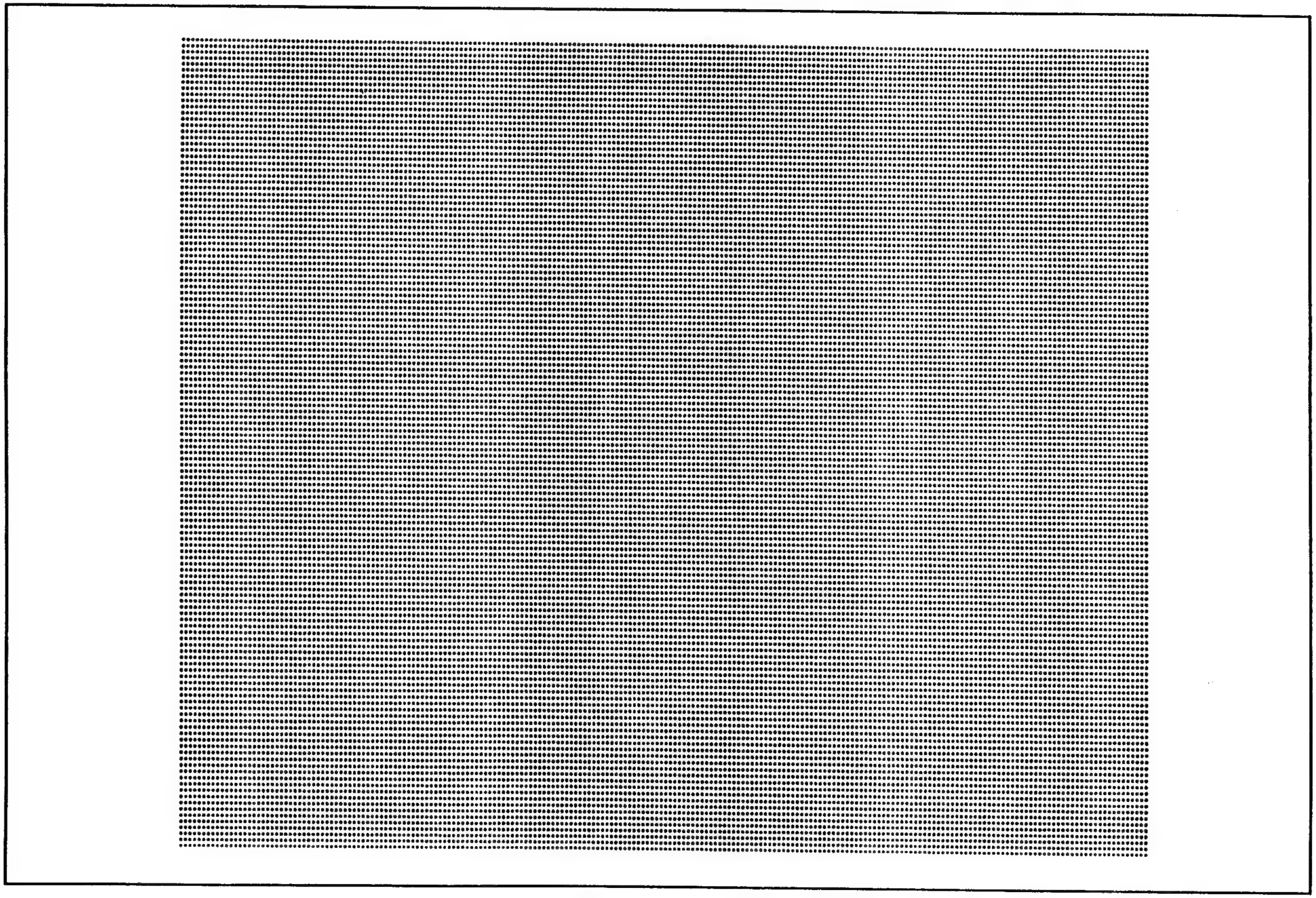


Figure A19. Beaverdam Creek DEM Data - No Filtering

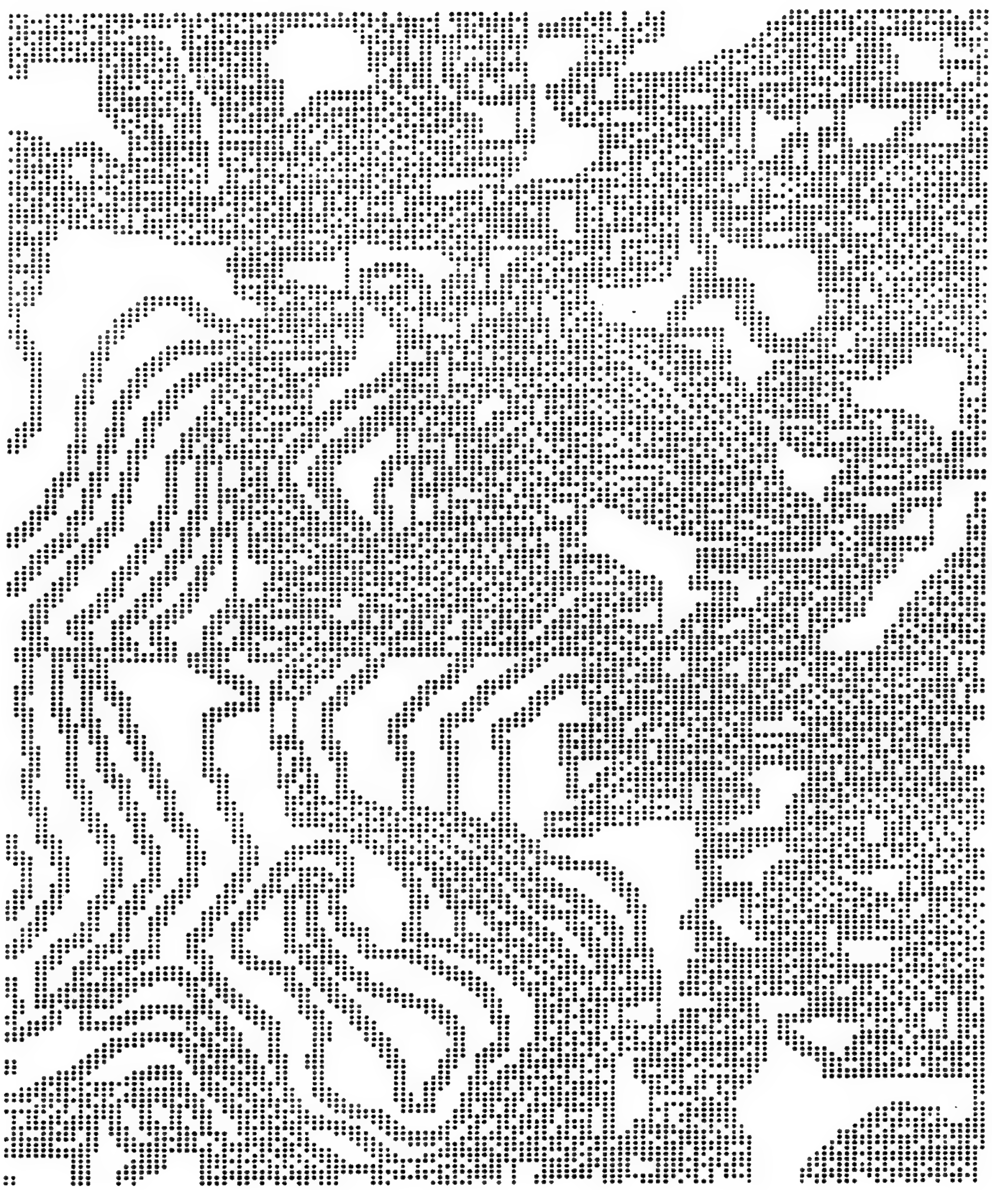


Figure A20. Beaverdam Creek DEM Data - Filter Threshold 0

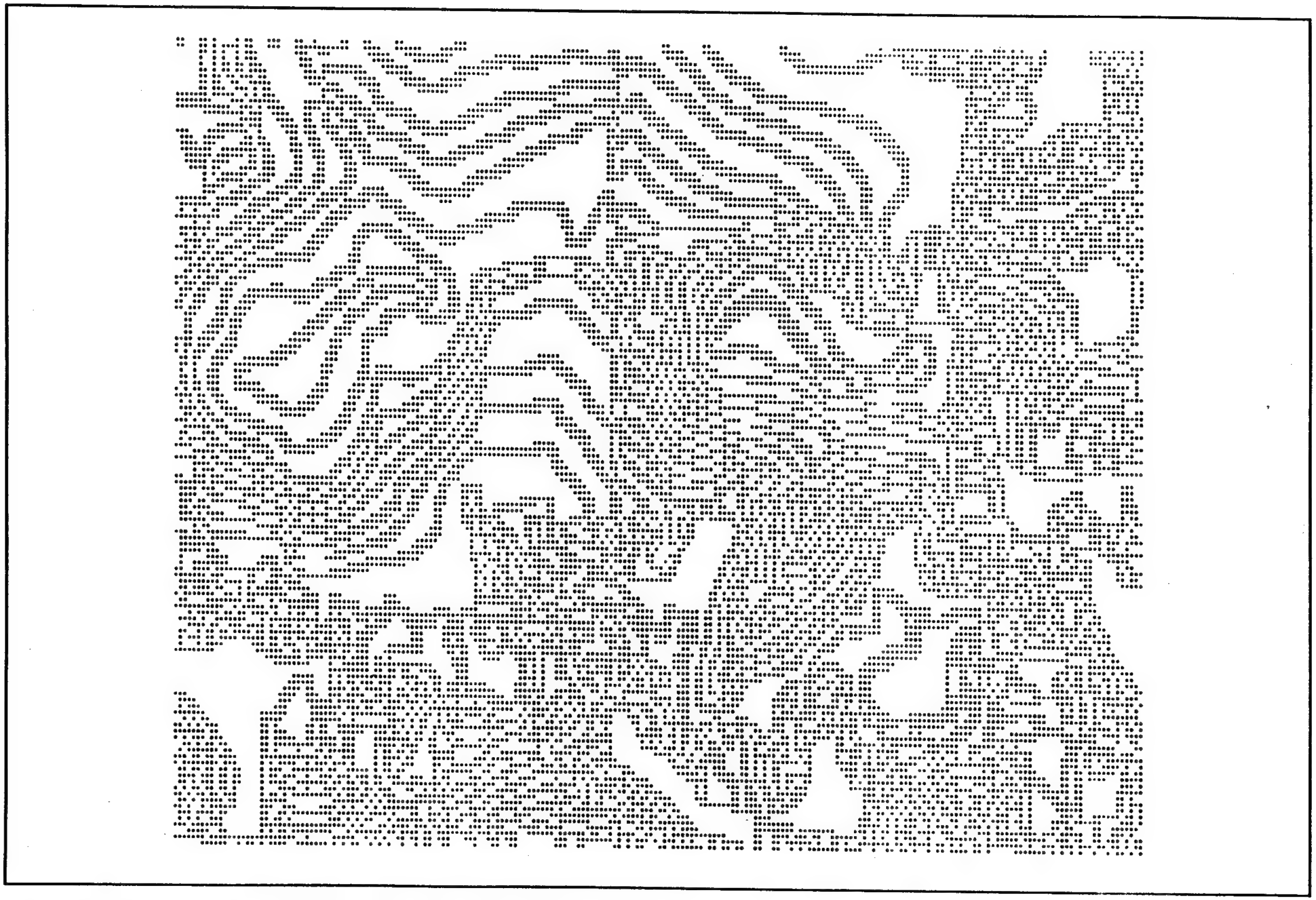


Figure A21. Beaverdam Creek DEM Data - Filter Threshold 1

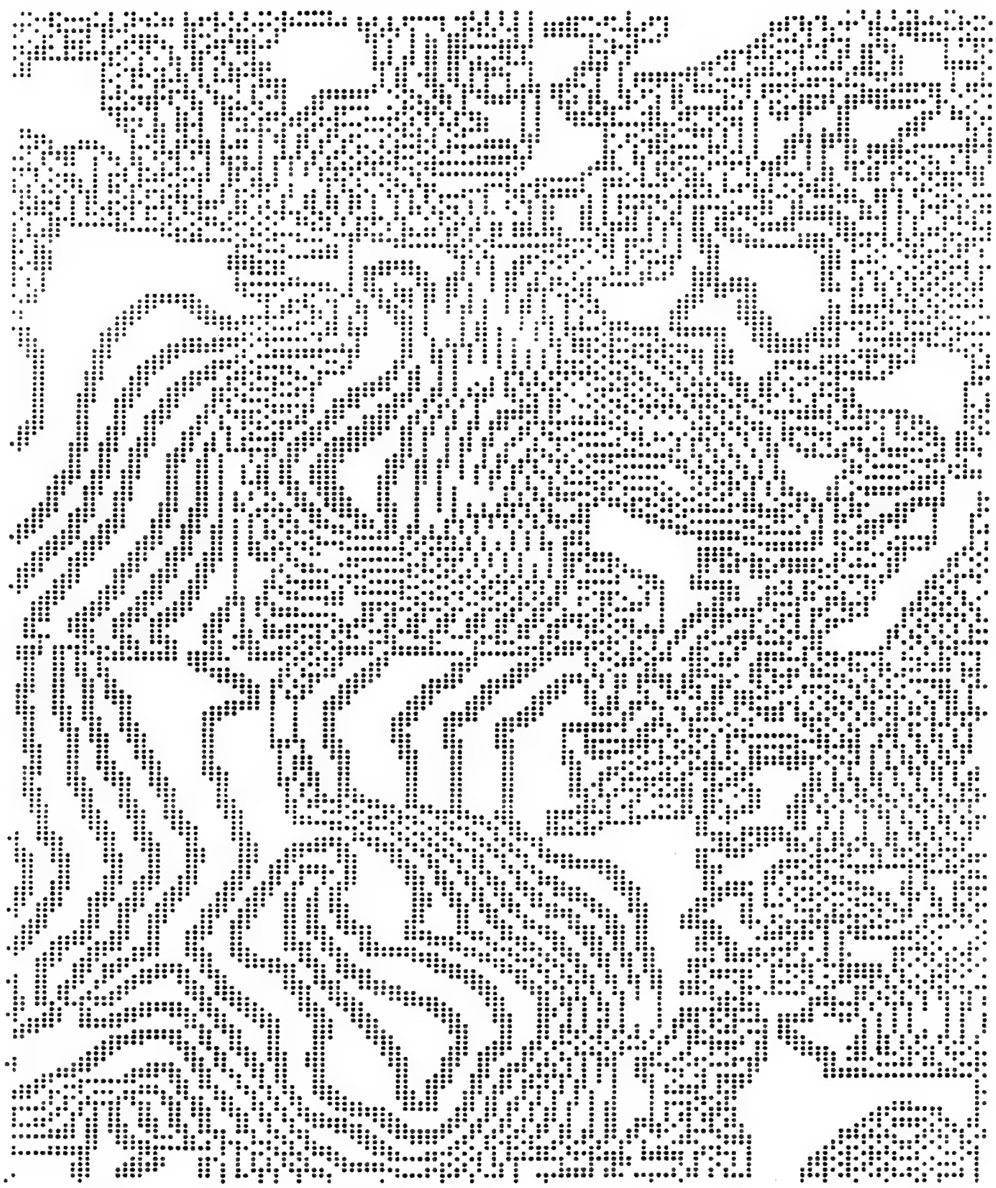


Figure A22. Beaverdam Creek DEM Data - Filter Threshold 2

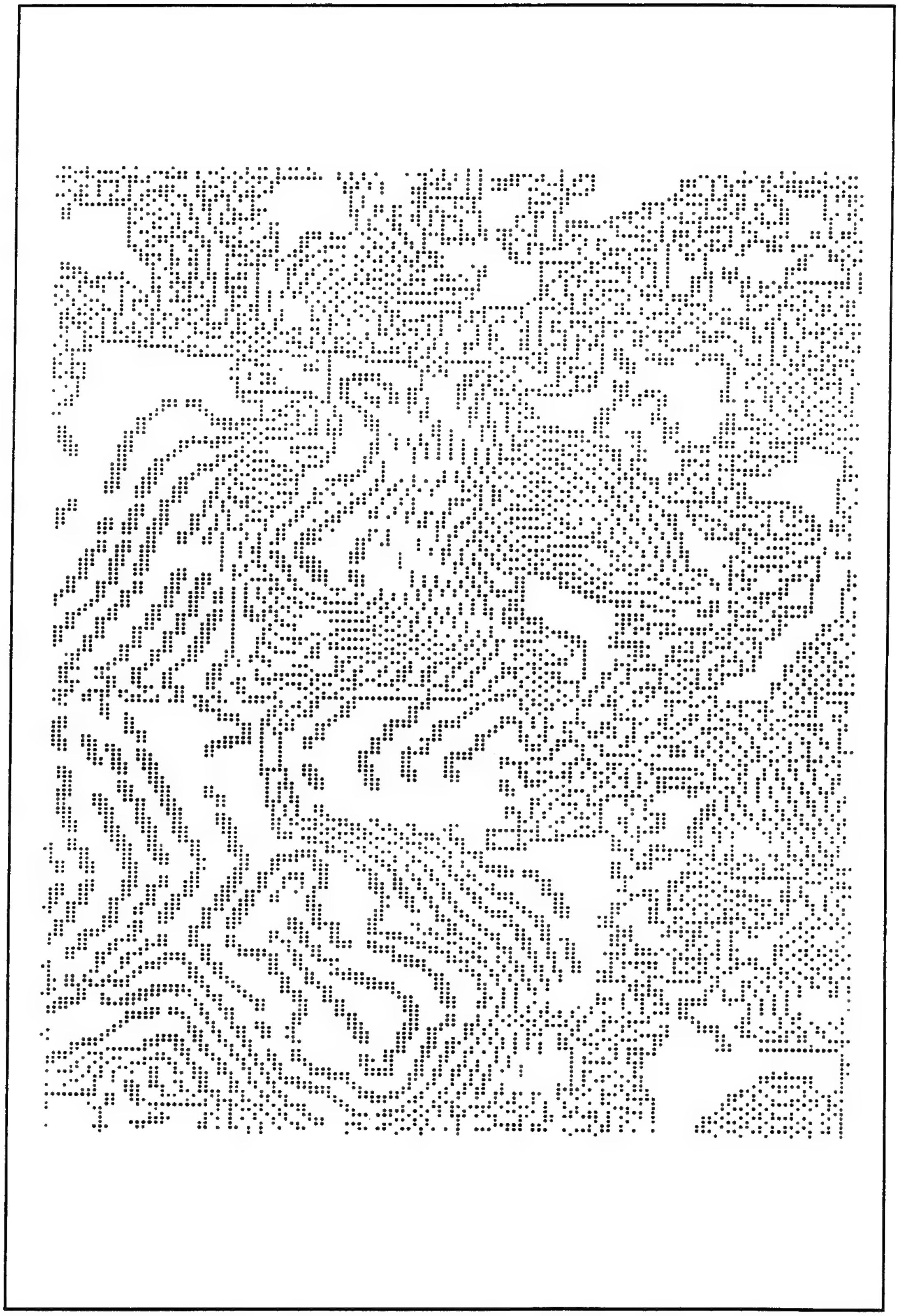


Figure A23. Beaverdam Creek DEM Data - Filter Threshold 3

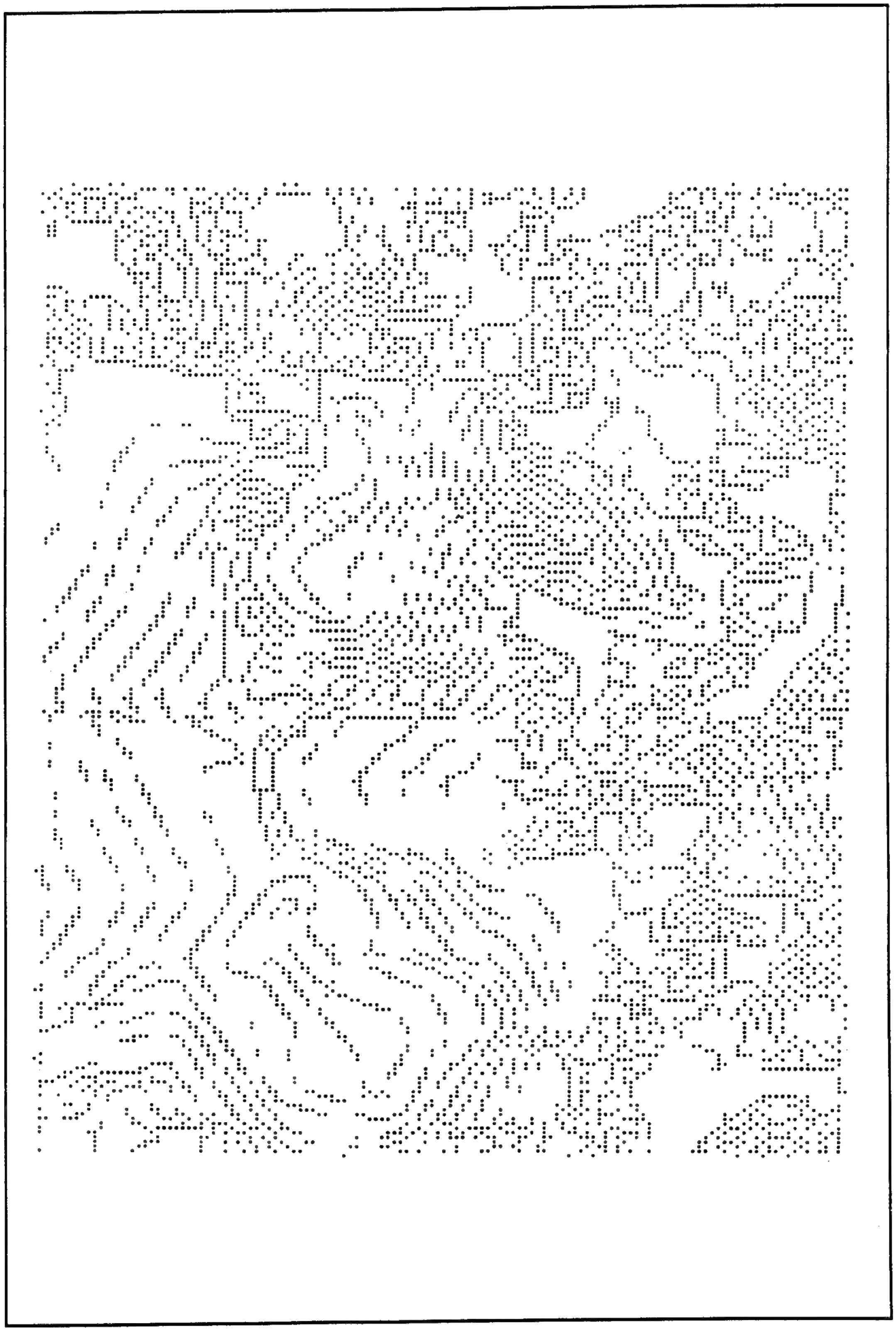


Figure A24.  Beaverdam Creek DEM Data - Filter Threshold 4

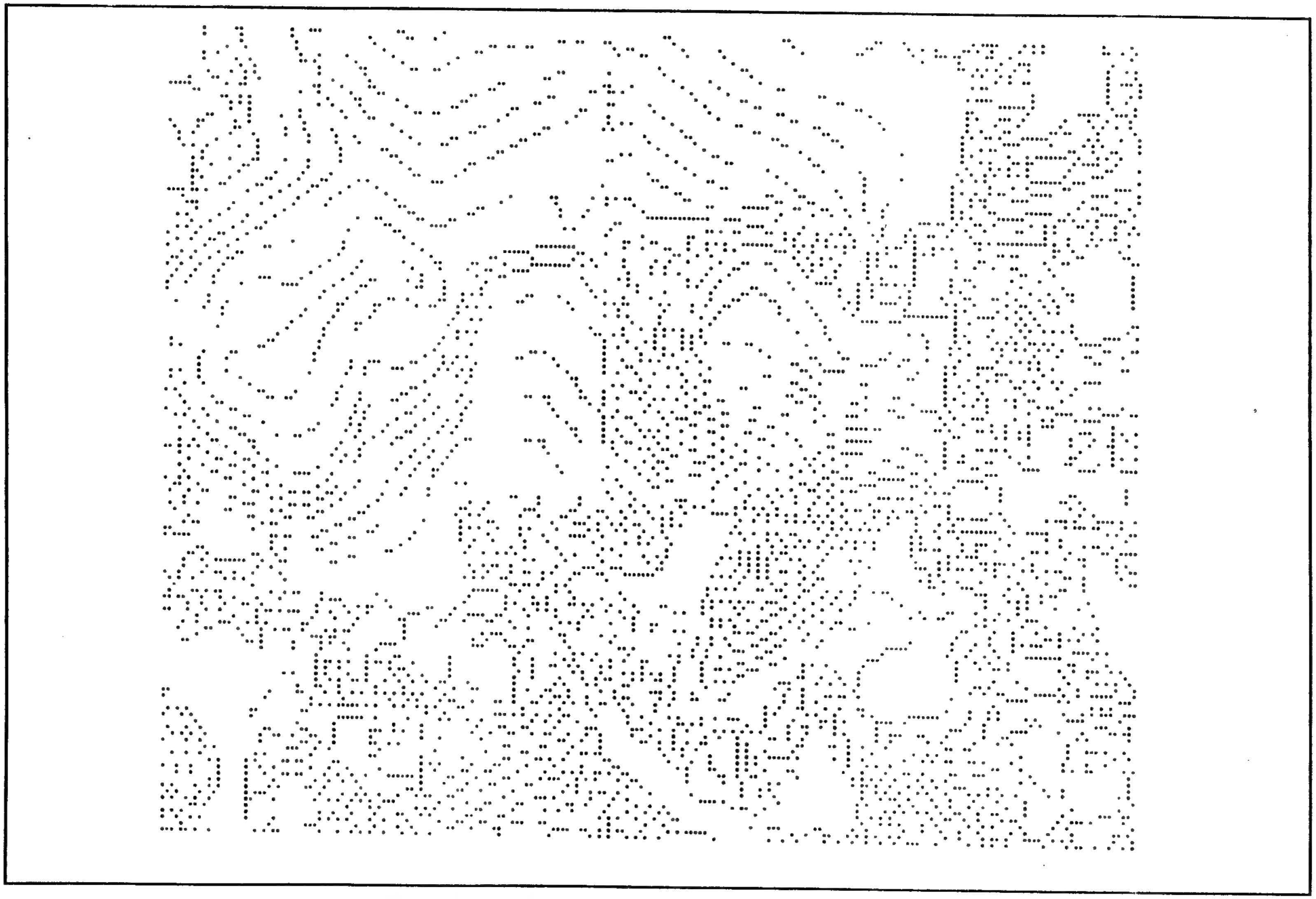


Figure A25. Beaverdam Creek DEM Data - Filter Threshold 5

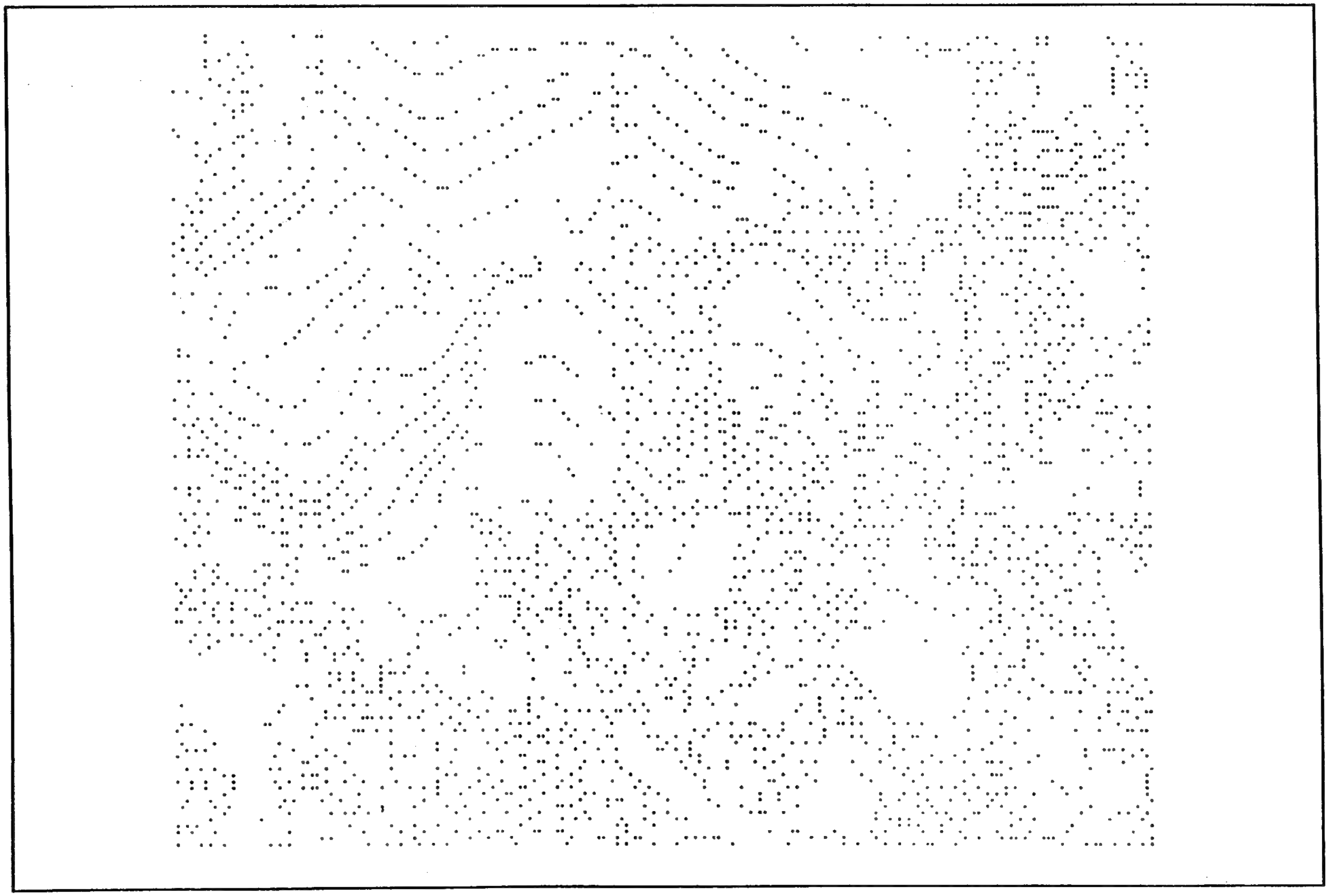


Figure A26.  Beaverdam Creek DEM Data - Filter Threshold 6

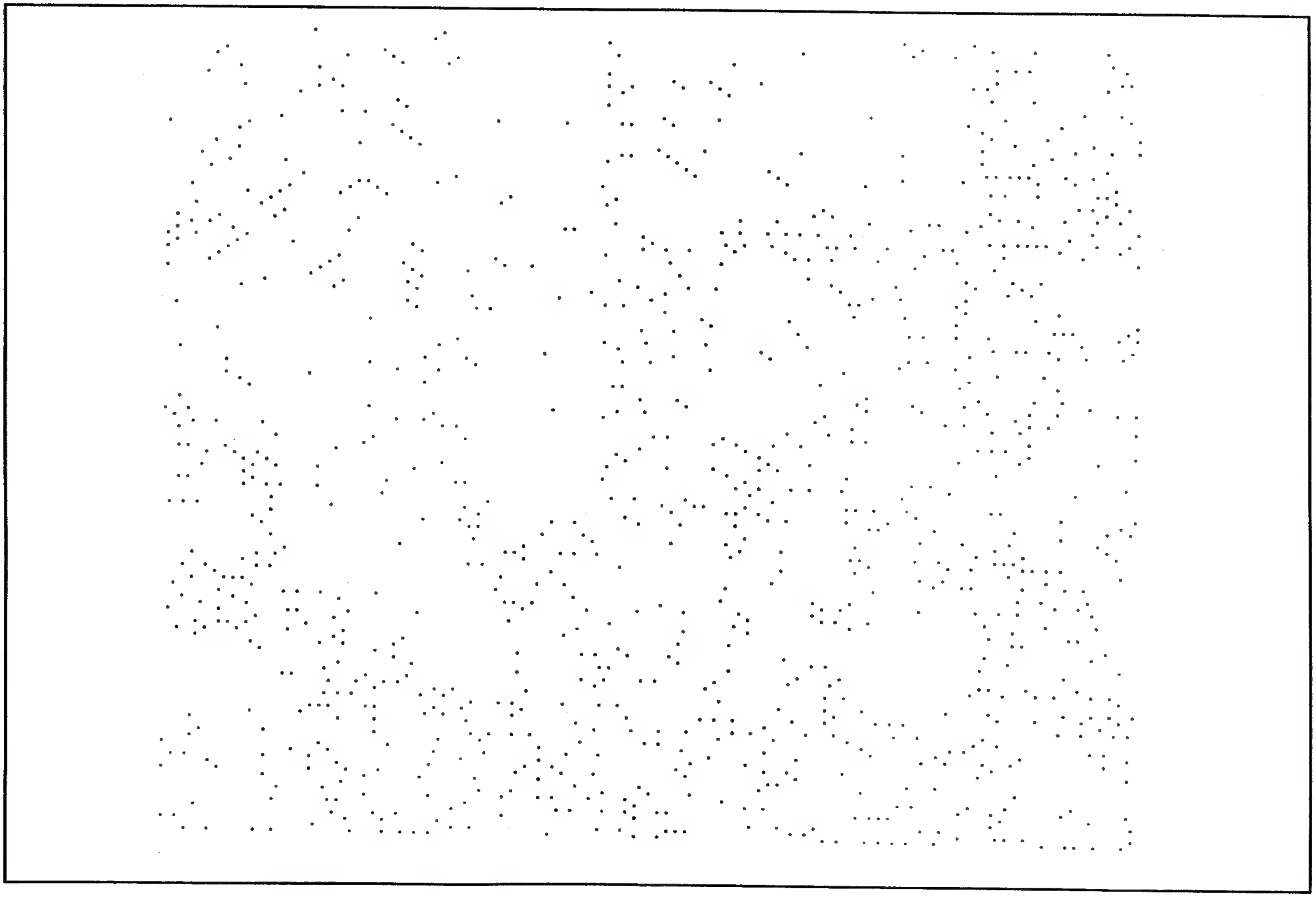


Figure A27. Beaverdam Creek DEM Data - Filter Threshold 7

# APPENDIX B

# PLOTS OF TINS FROM FILTERED ELEVATION DATA

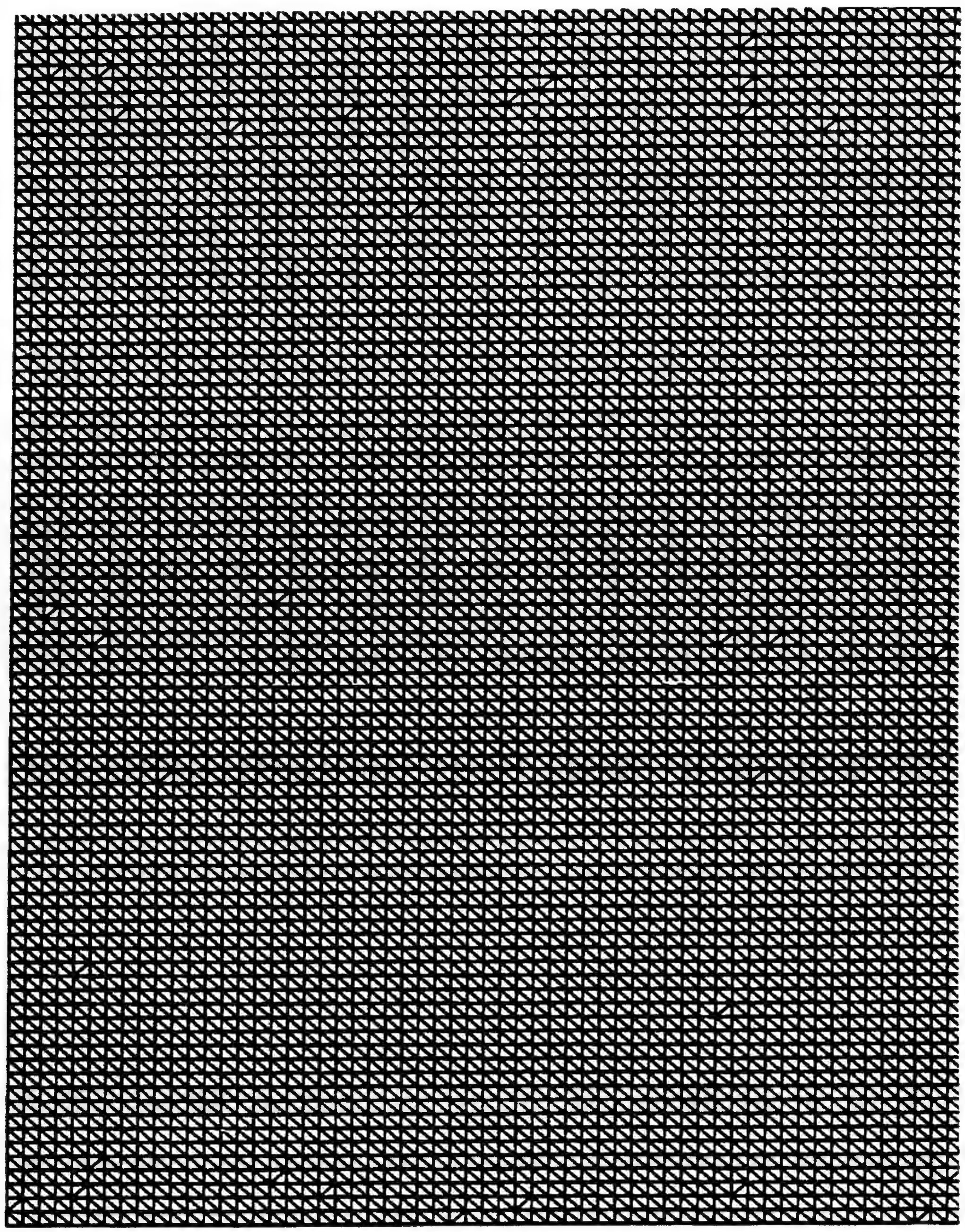


Figure B1. Rock Creek TIN - No Filtering

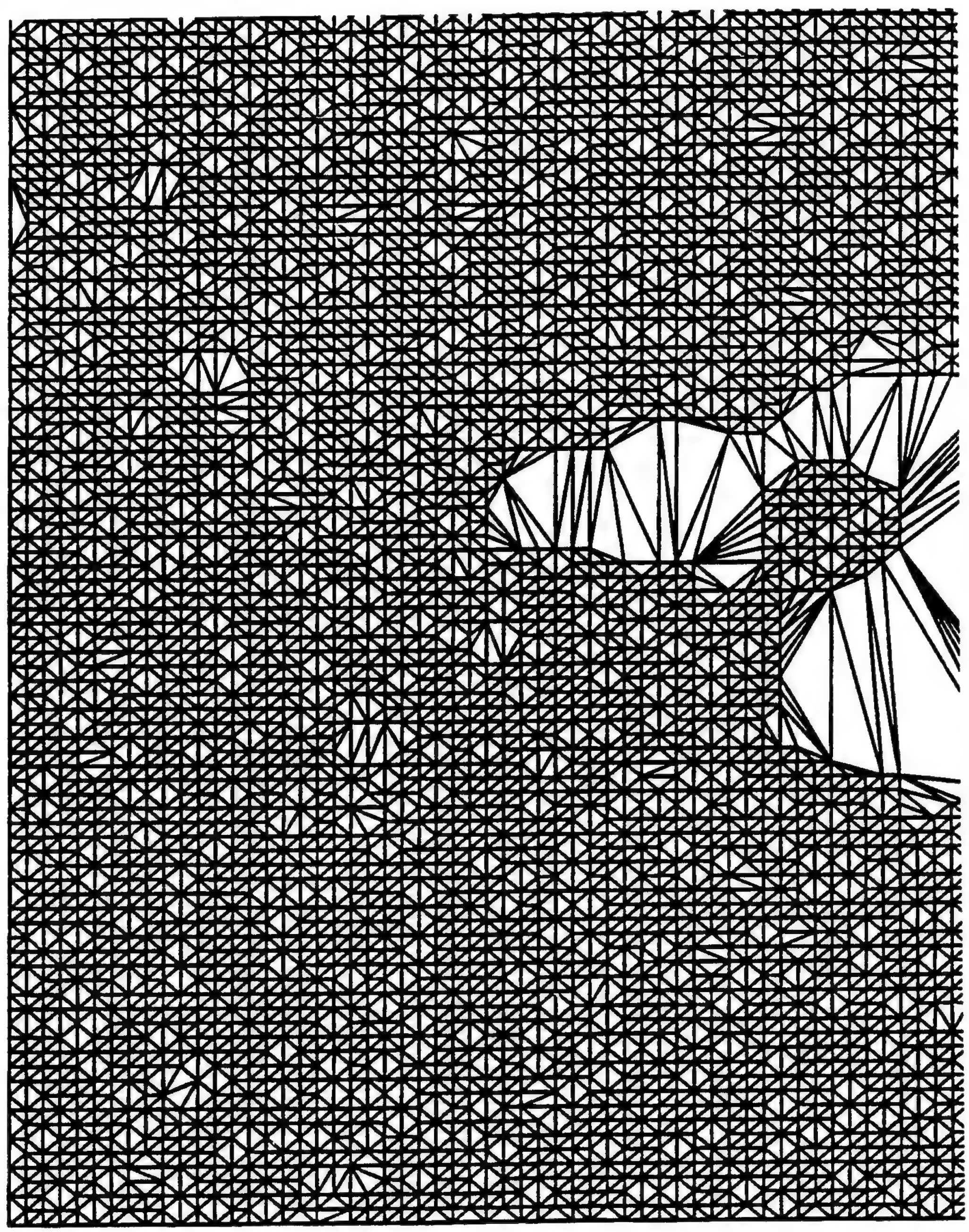


Figure B2. Rock Creek TIN - Filter Threshold 0

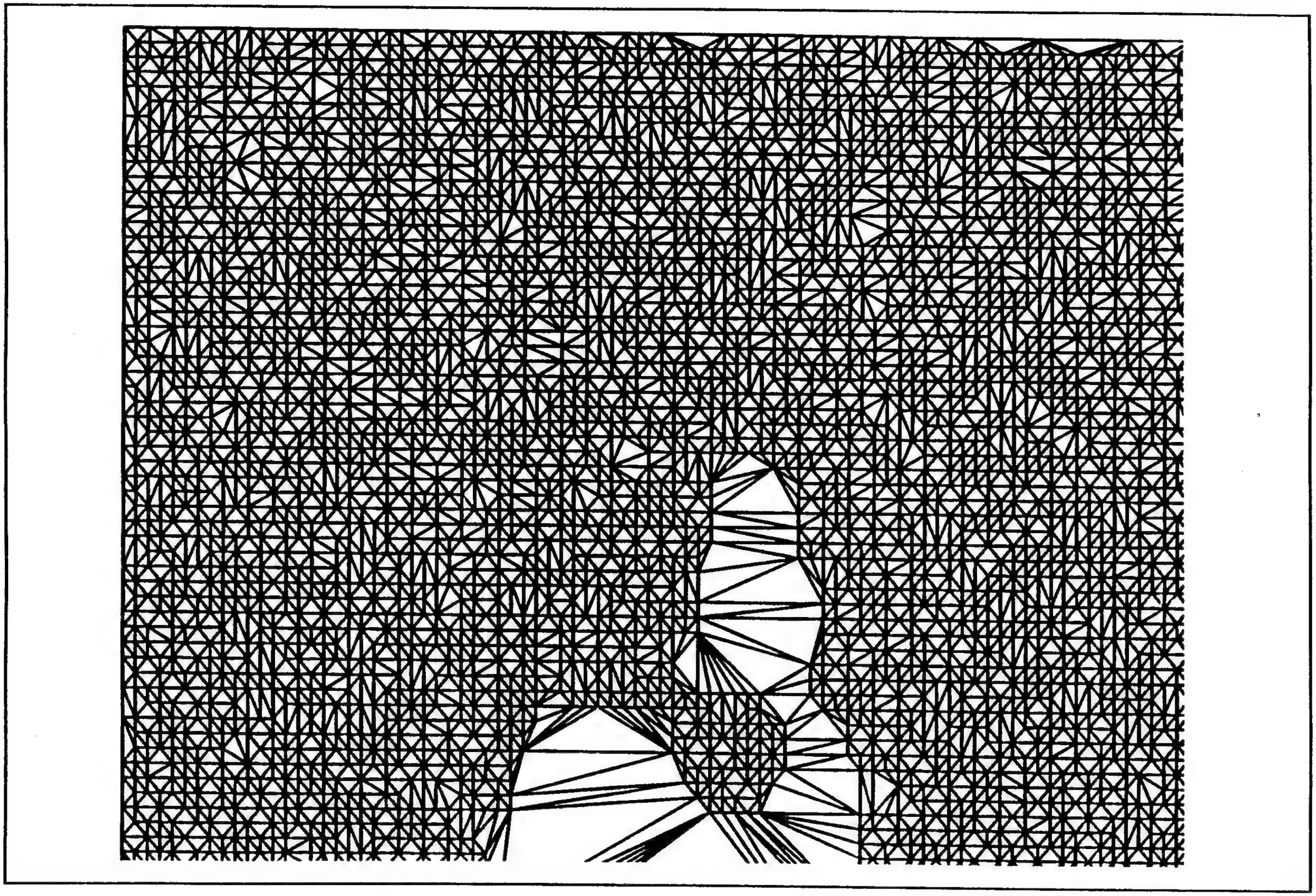


Figure B3. Rock Creek TIN - Filter Threshold 1

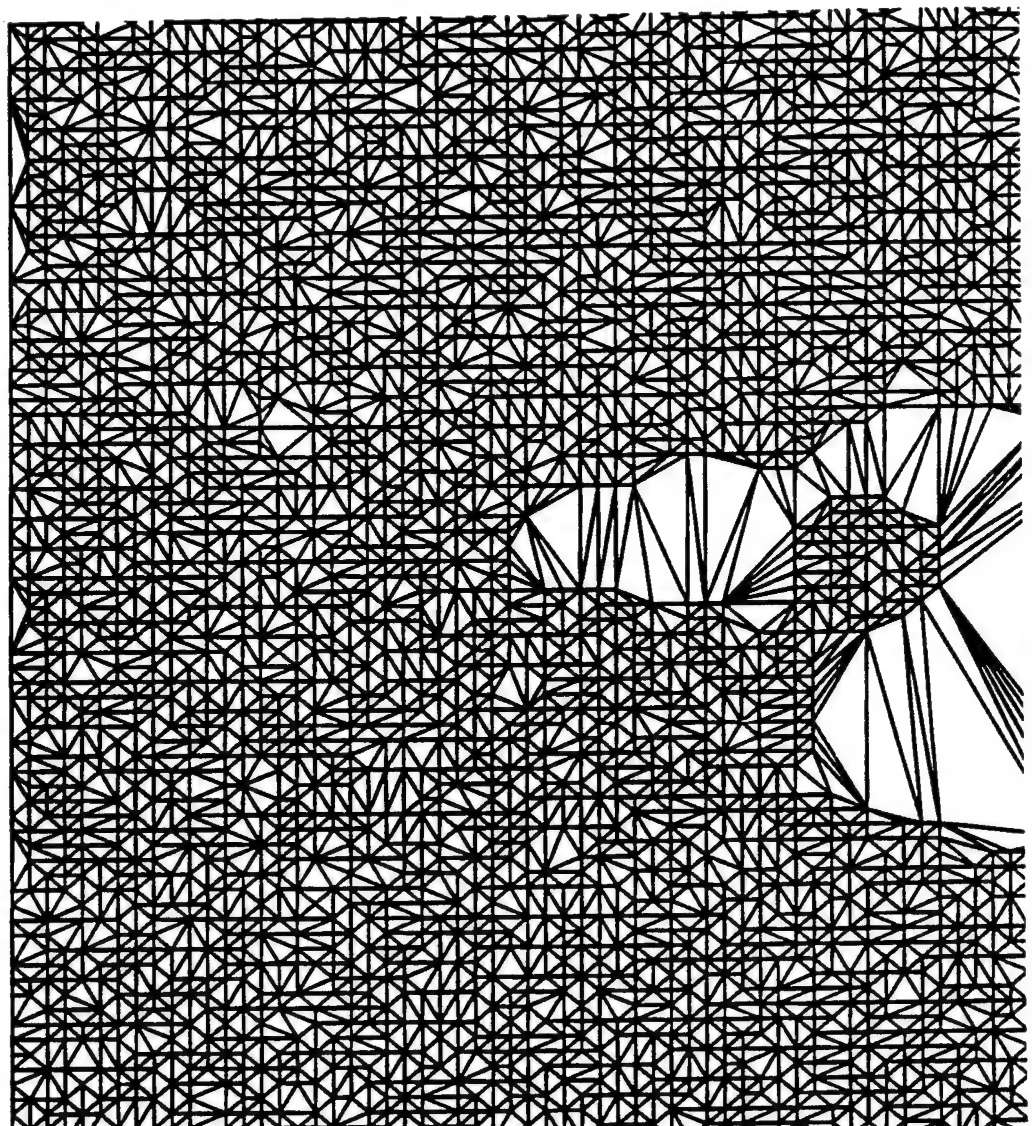


Figure B4. Rock Creek TIN - Filter Threshold 2

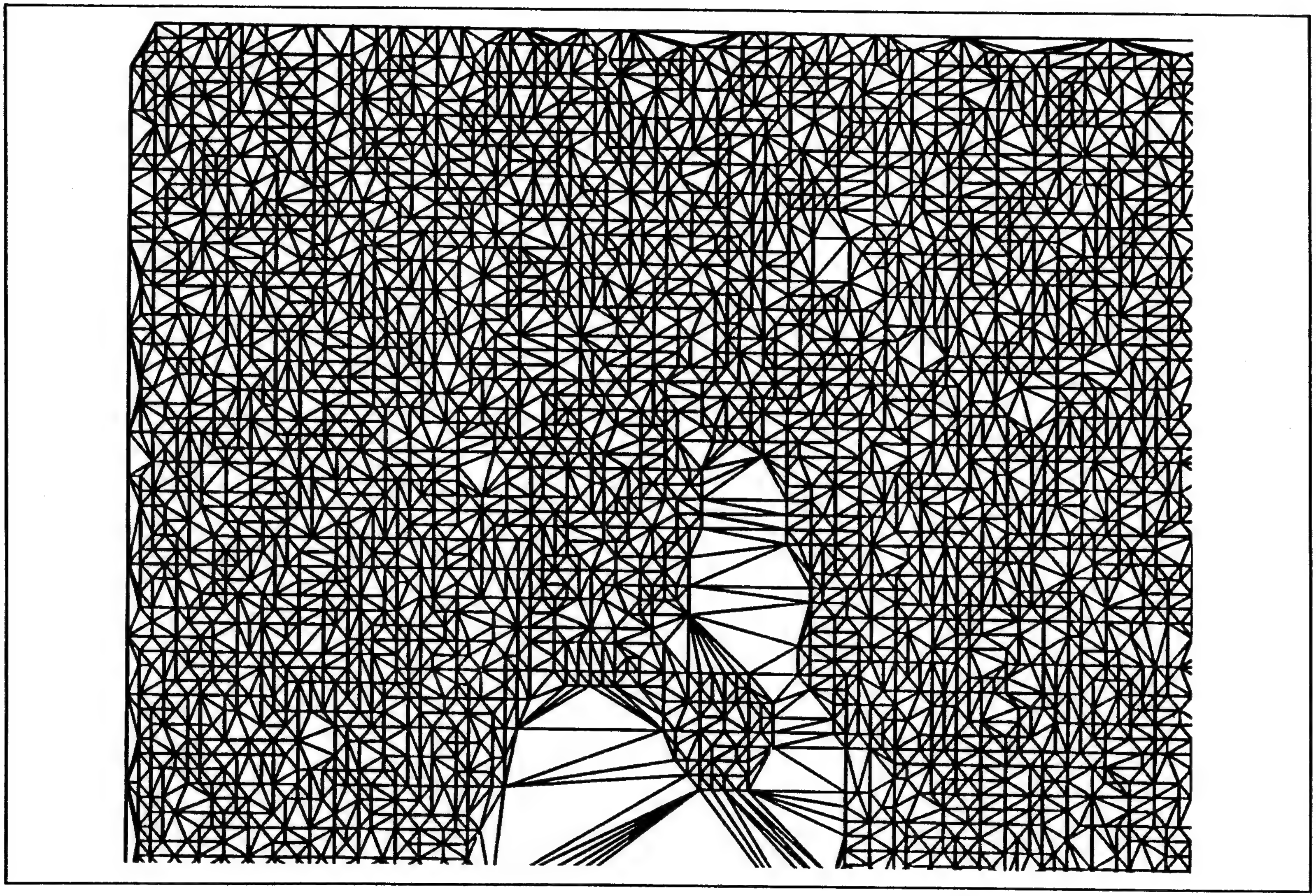


Figure B5.  Rock Creek TIN - Filter Threshold 3

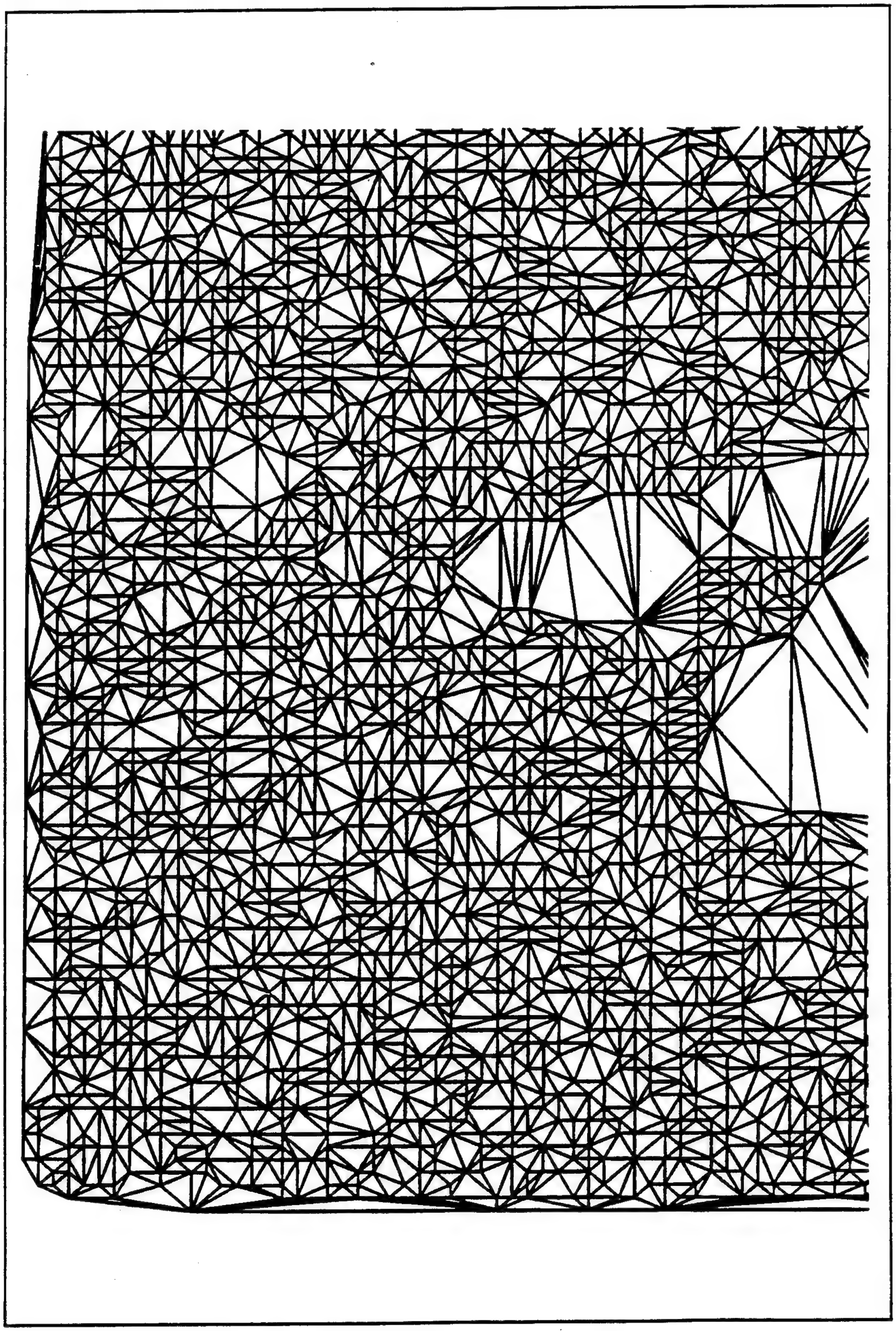


Figure B6. Rock Creek TIN - Filter Threshold 4

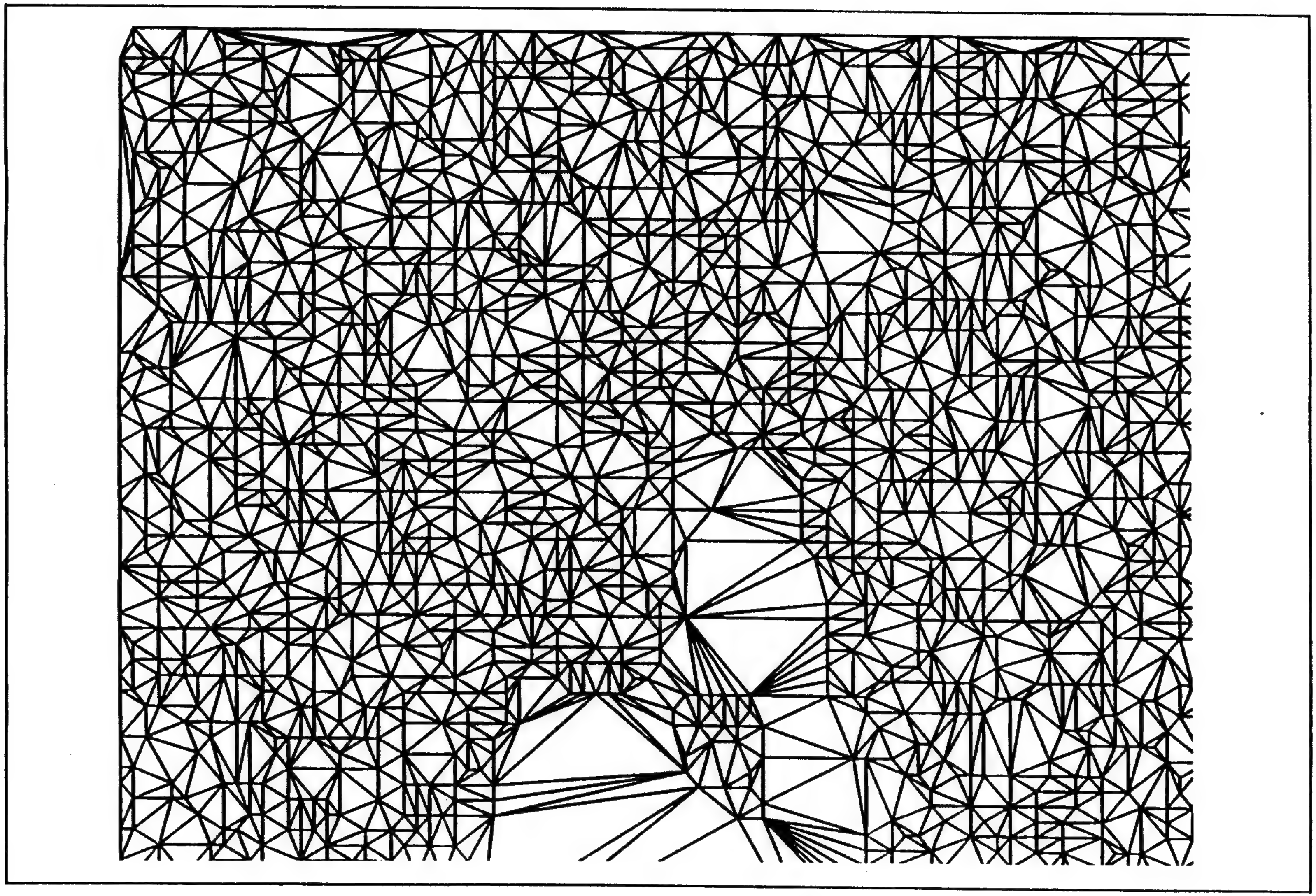


Figure B7. Rock Creek TIN - Filter Threshold 5

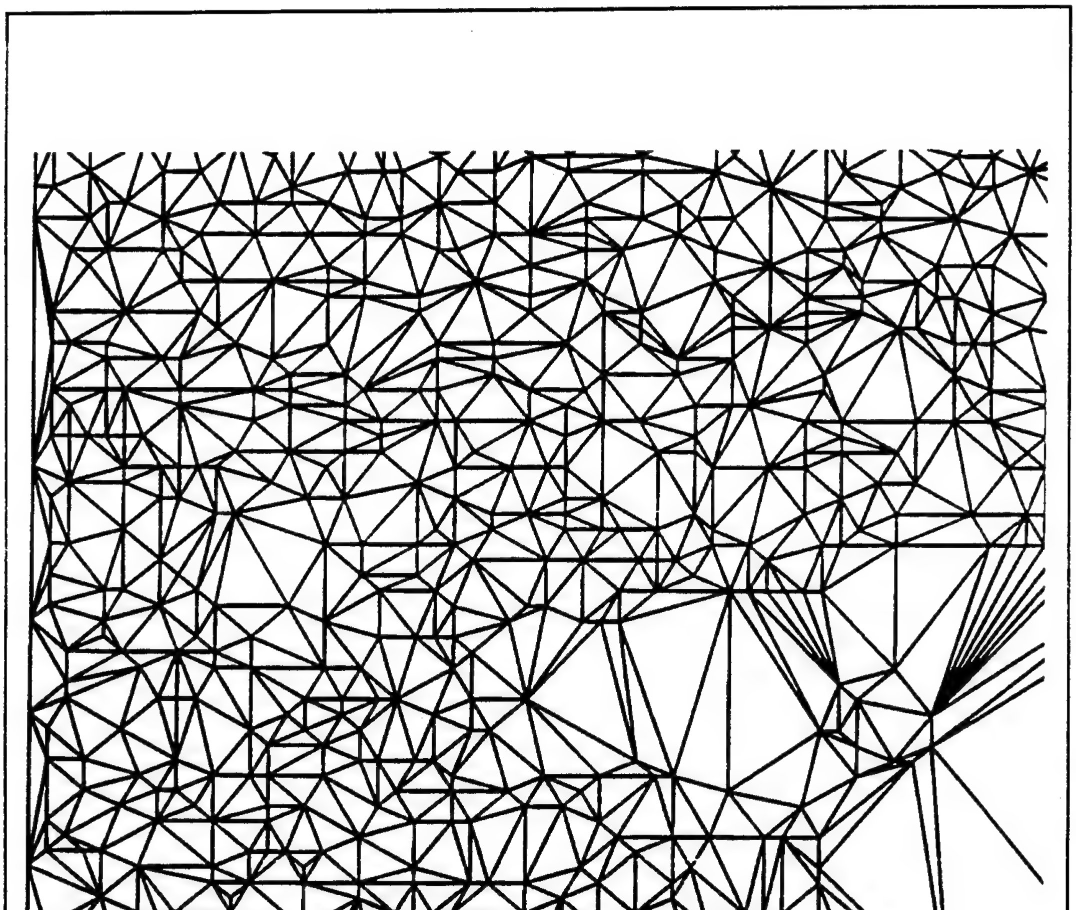


Figure B8. Rock Creek TIN - Filter Threshold 6

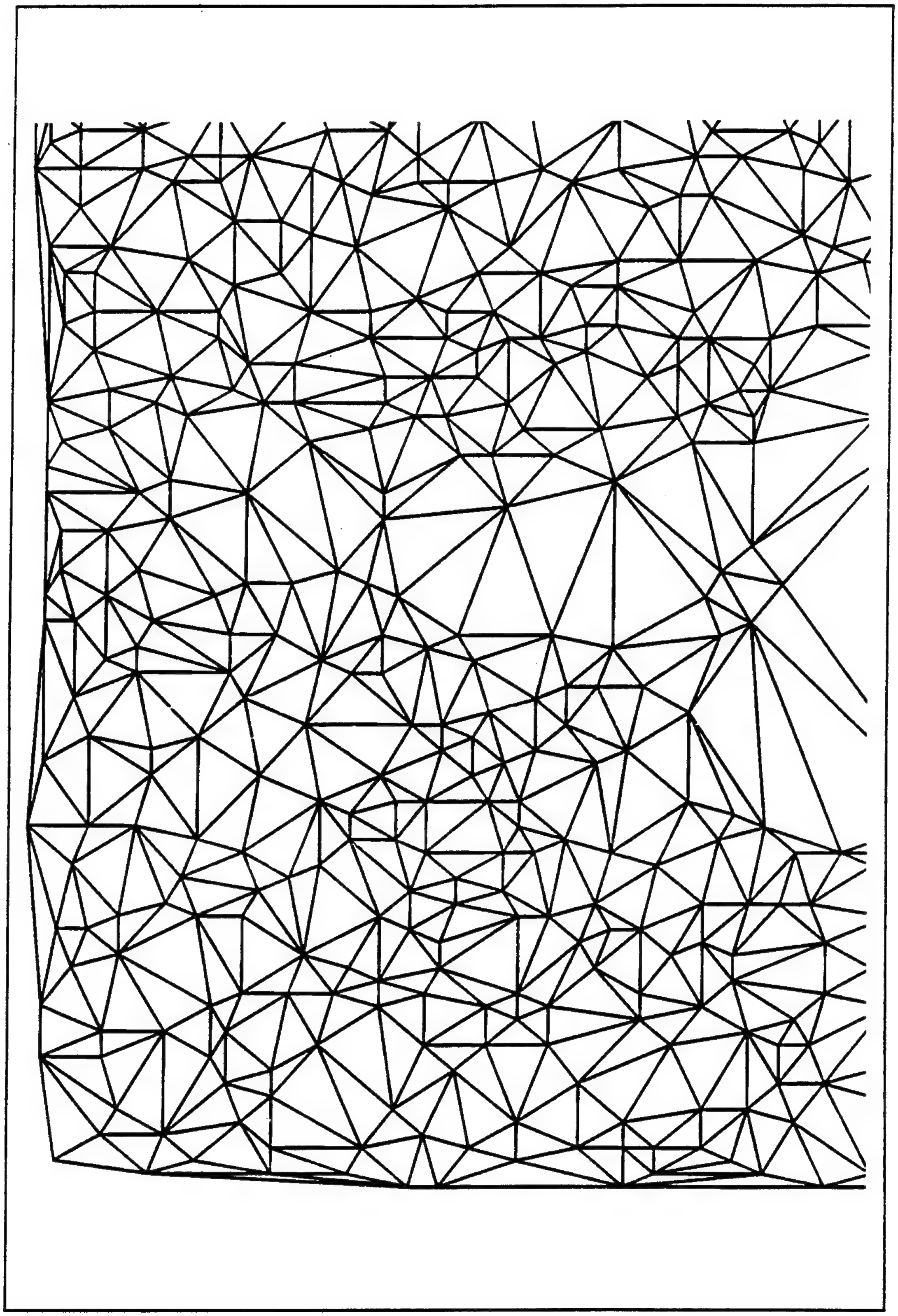


Figure B9. Rock Creek TIN - Filter Threshold 7

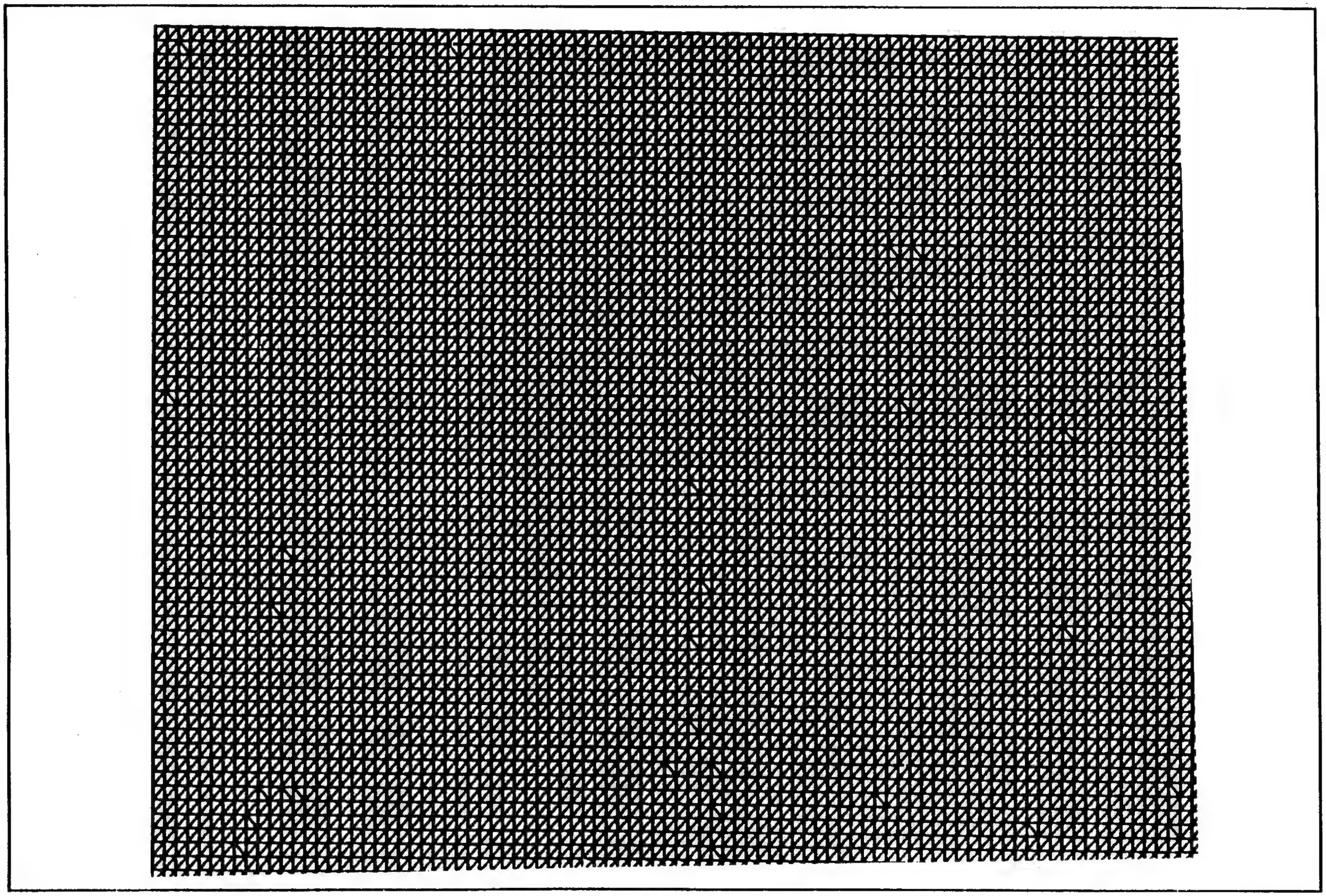


Figure B10. Long Creek TIN - No Filtering

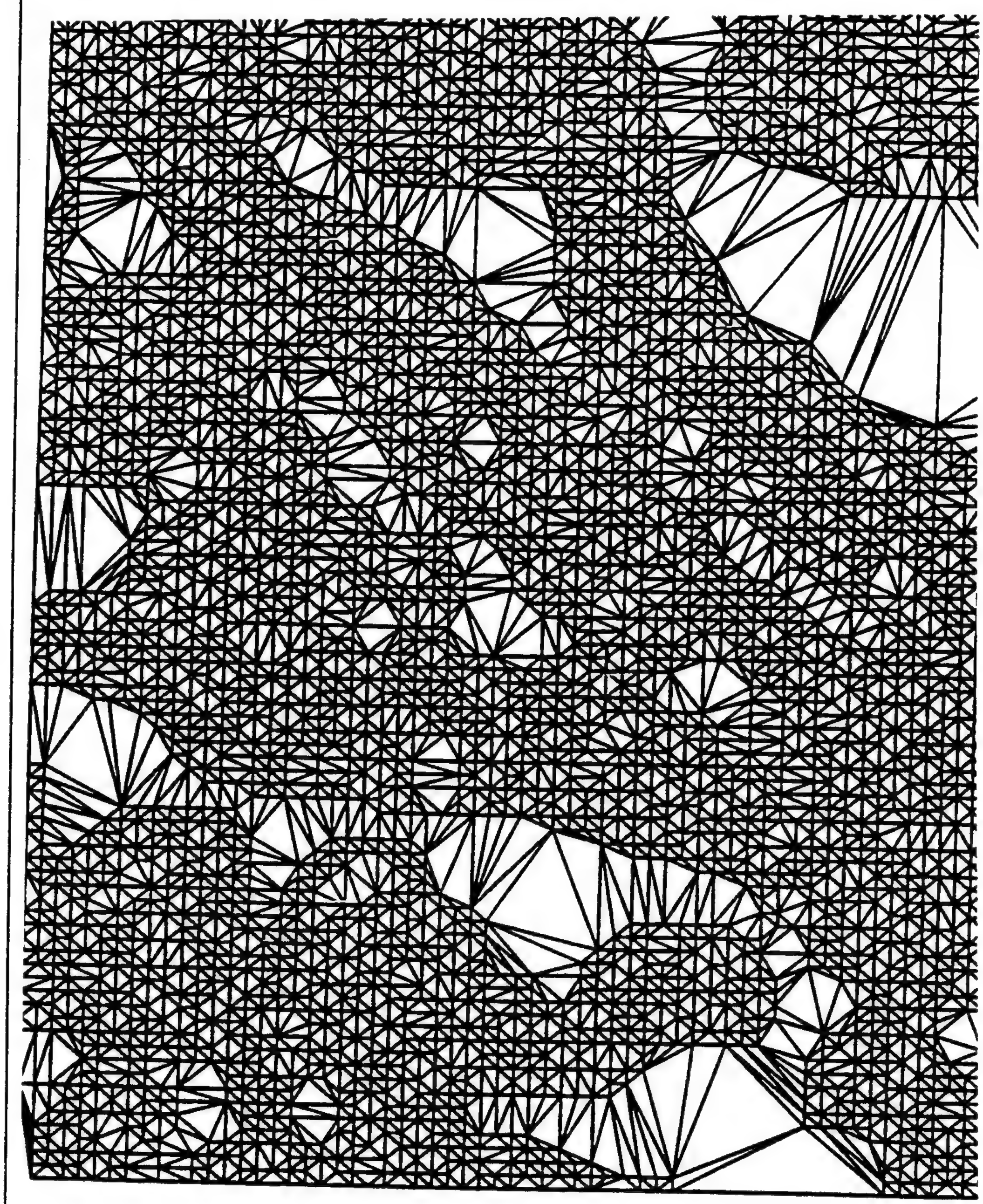


Figure B11. Long Creek TIN - Filter Threshold 0

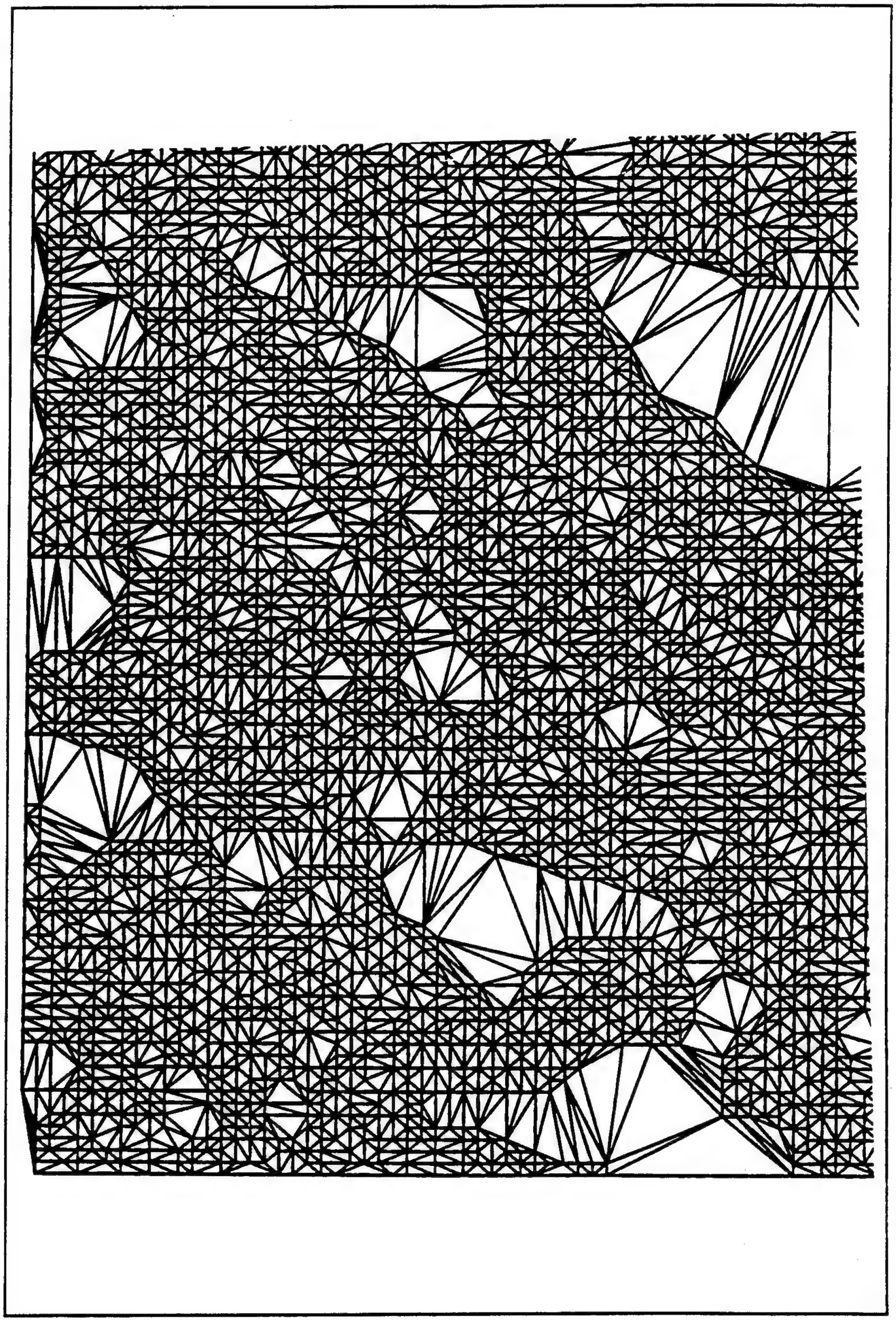


Figure B12. Long Creek TIN - Filter Threshold 1

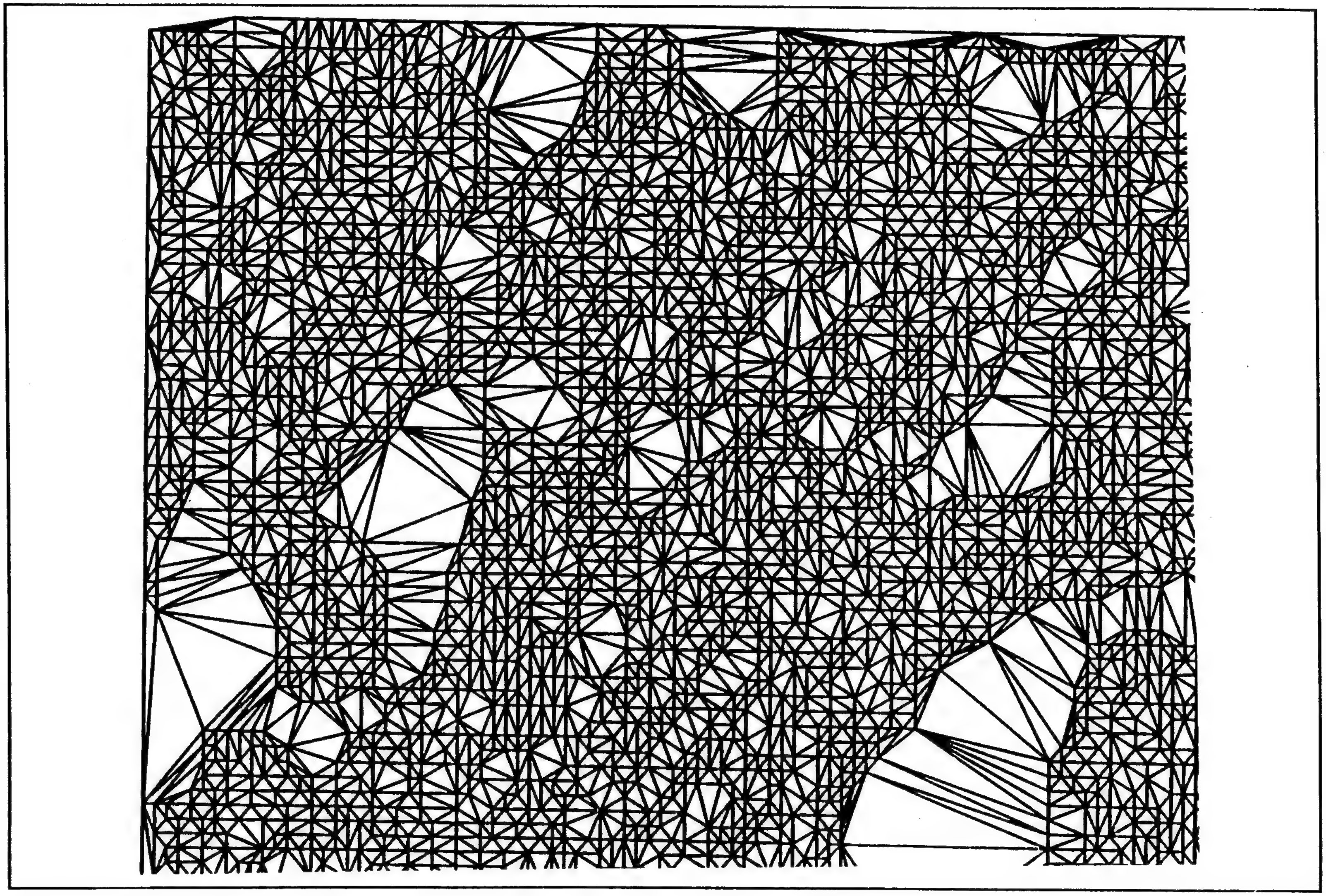


Figure B13. Long Creek TIN - Filter Threshold 2

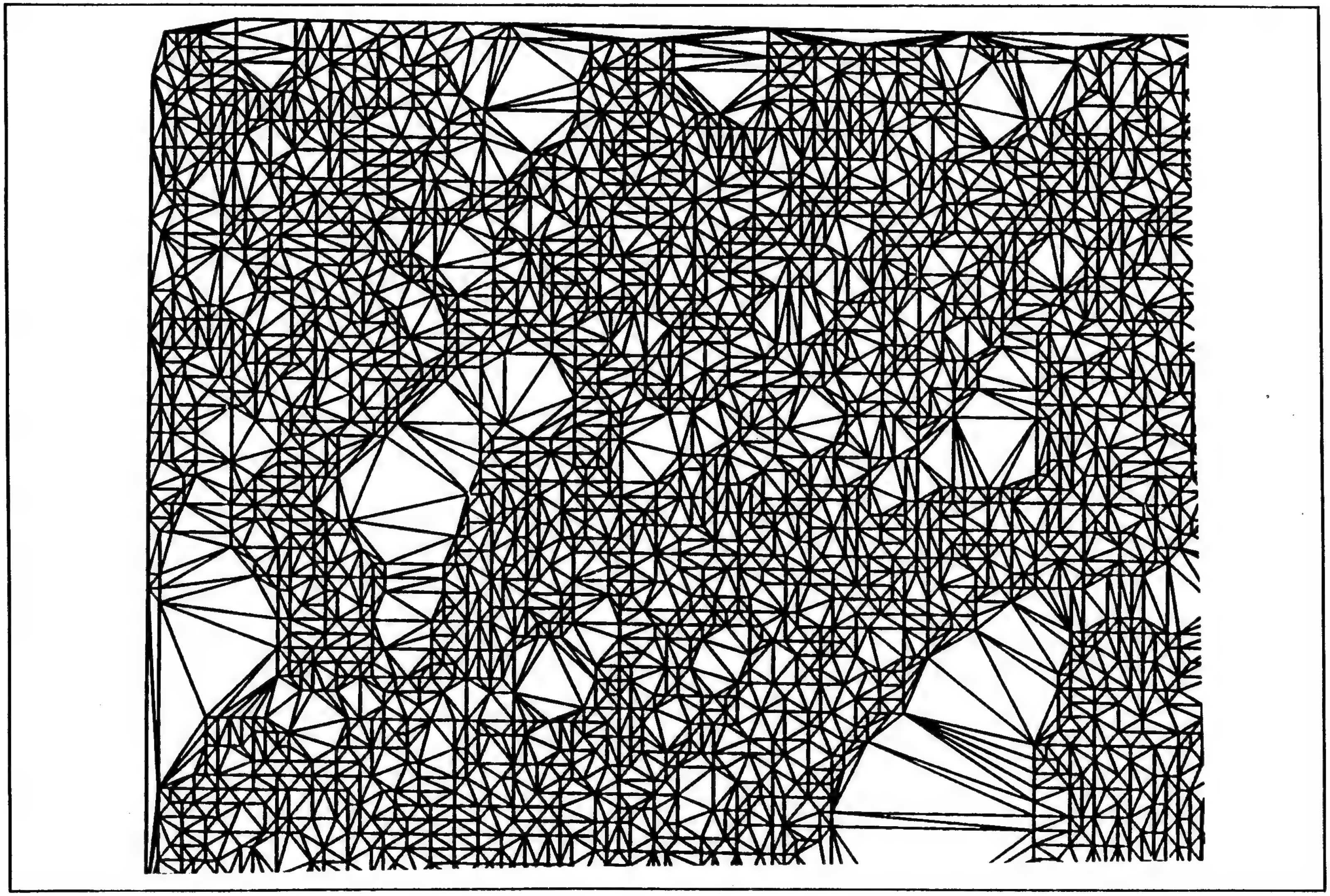


Figure B14.  Long Creek TIN - Filter Threshold 3

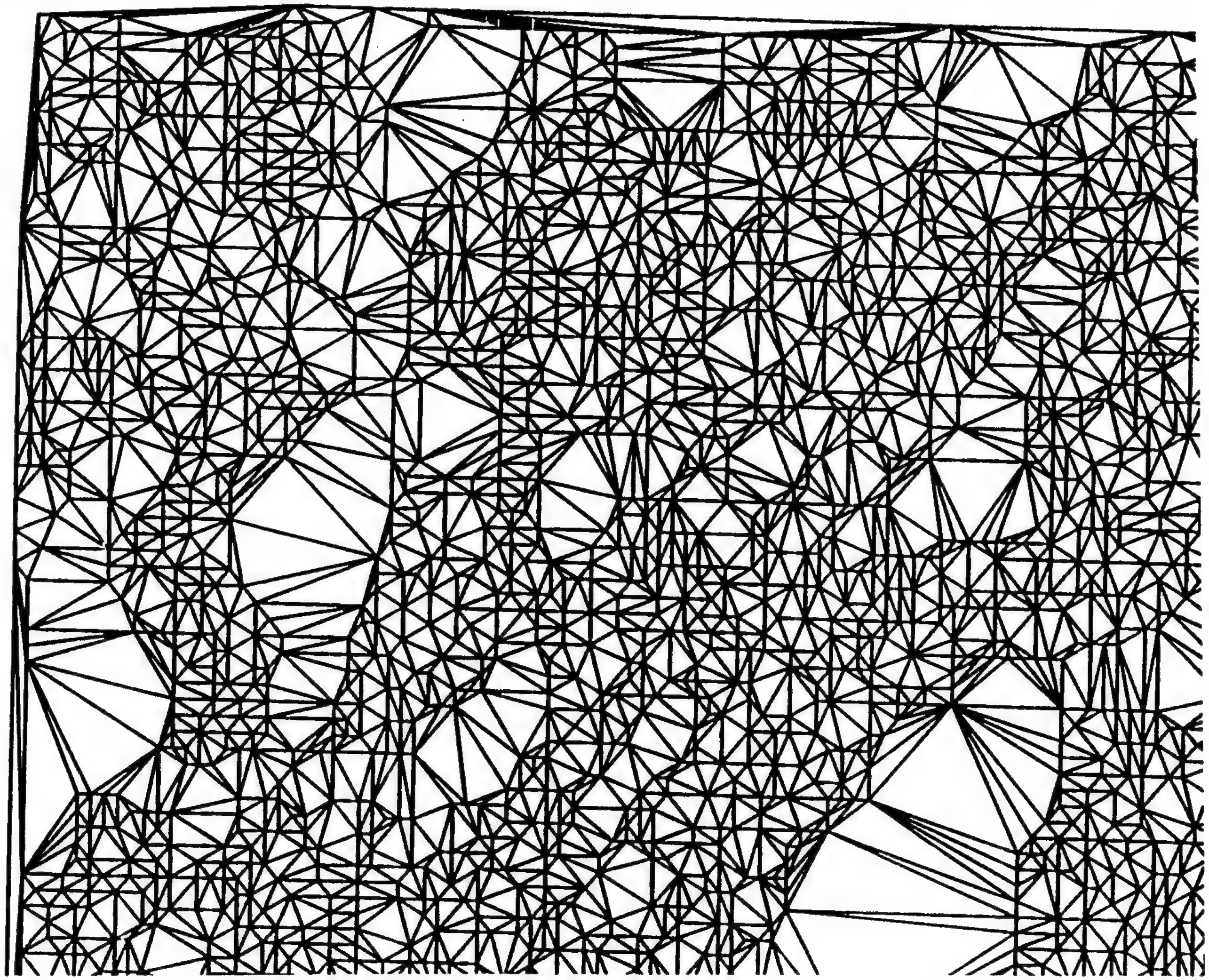


Figure B15. Long Creek TIN - Filter Threshold 4

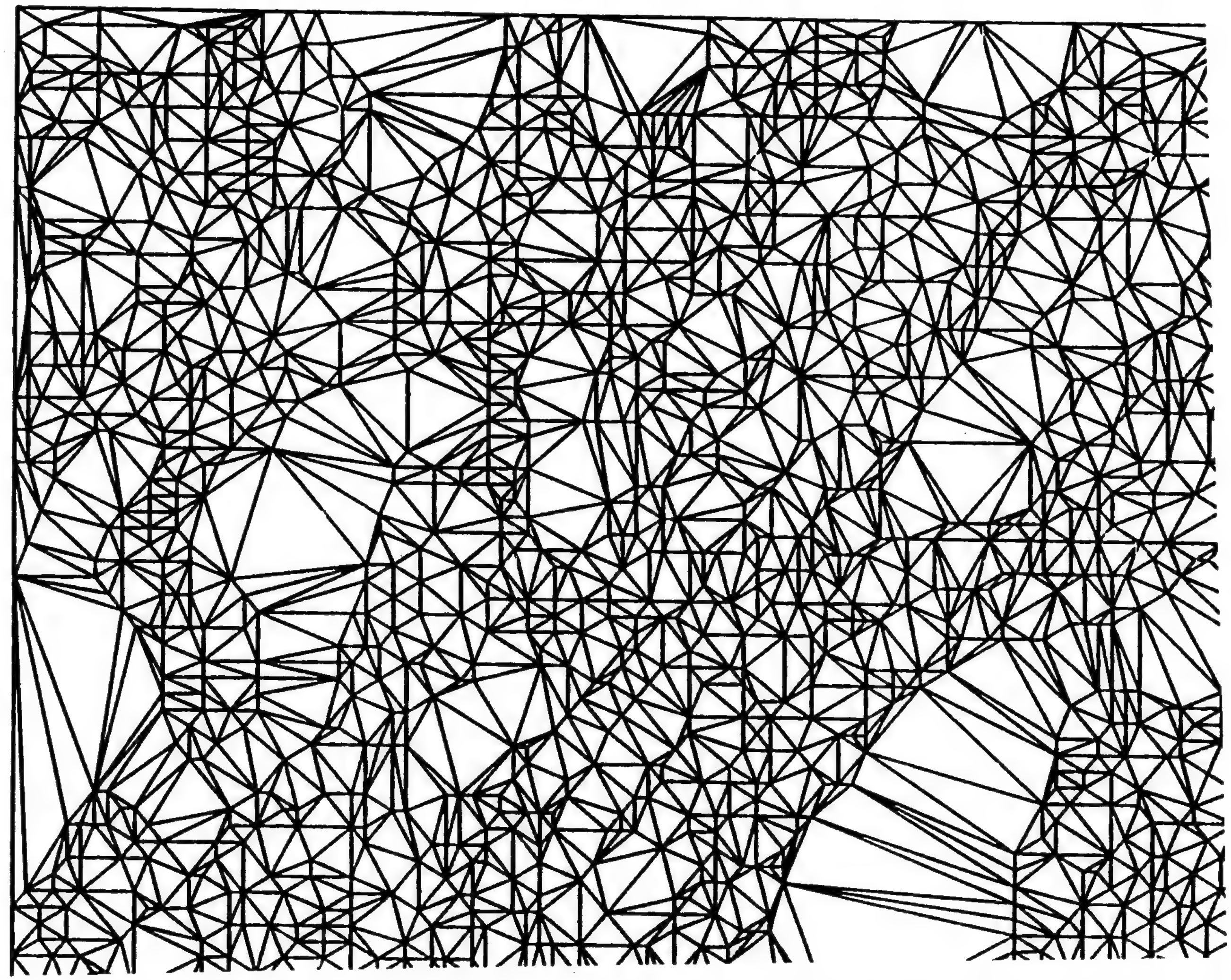


Figure B16.  Long Creek TIN - Filter Threshold 5

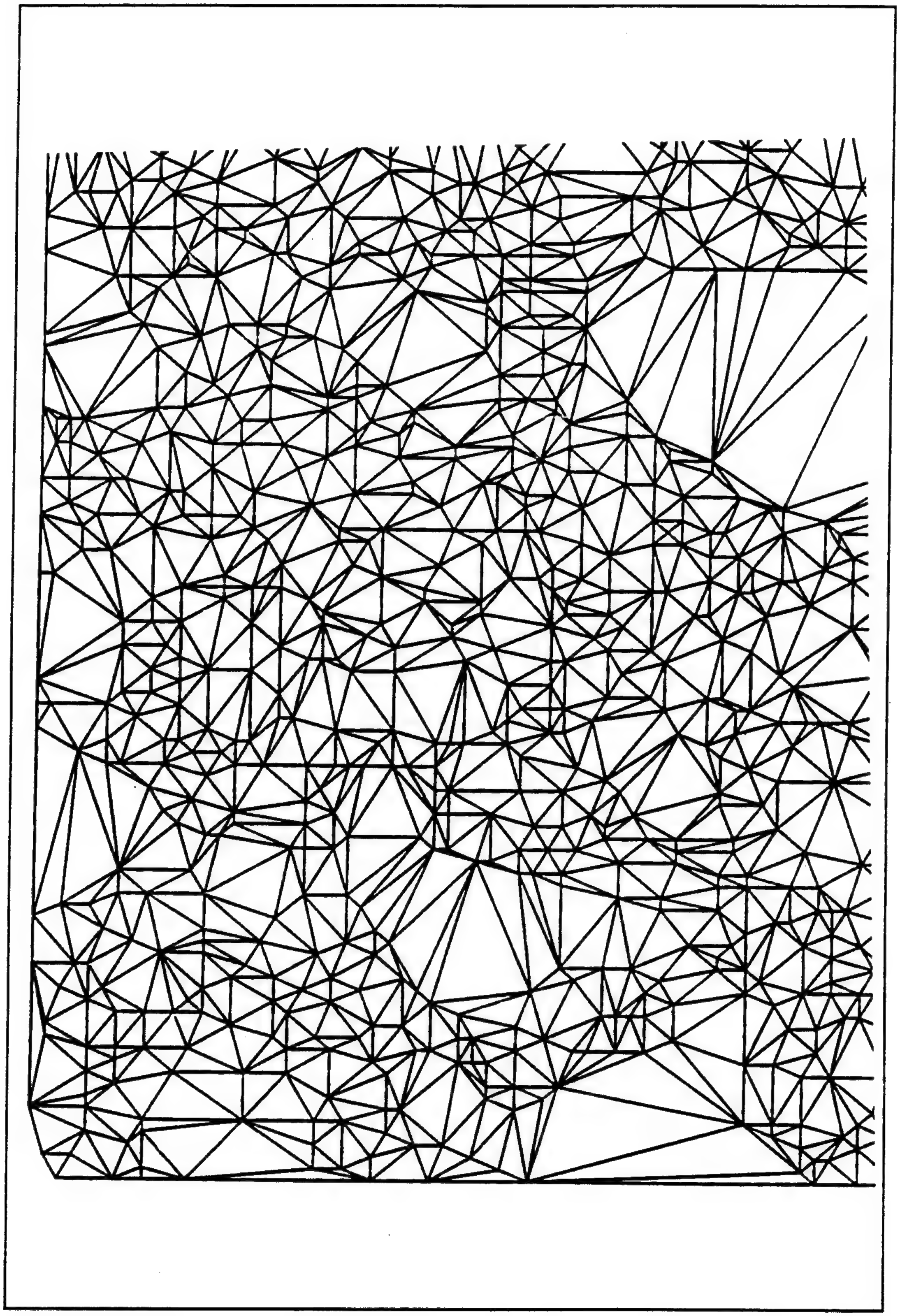


Figure B17. Long Creek TIN - Filter Threshold 6

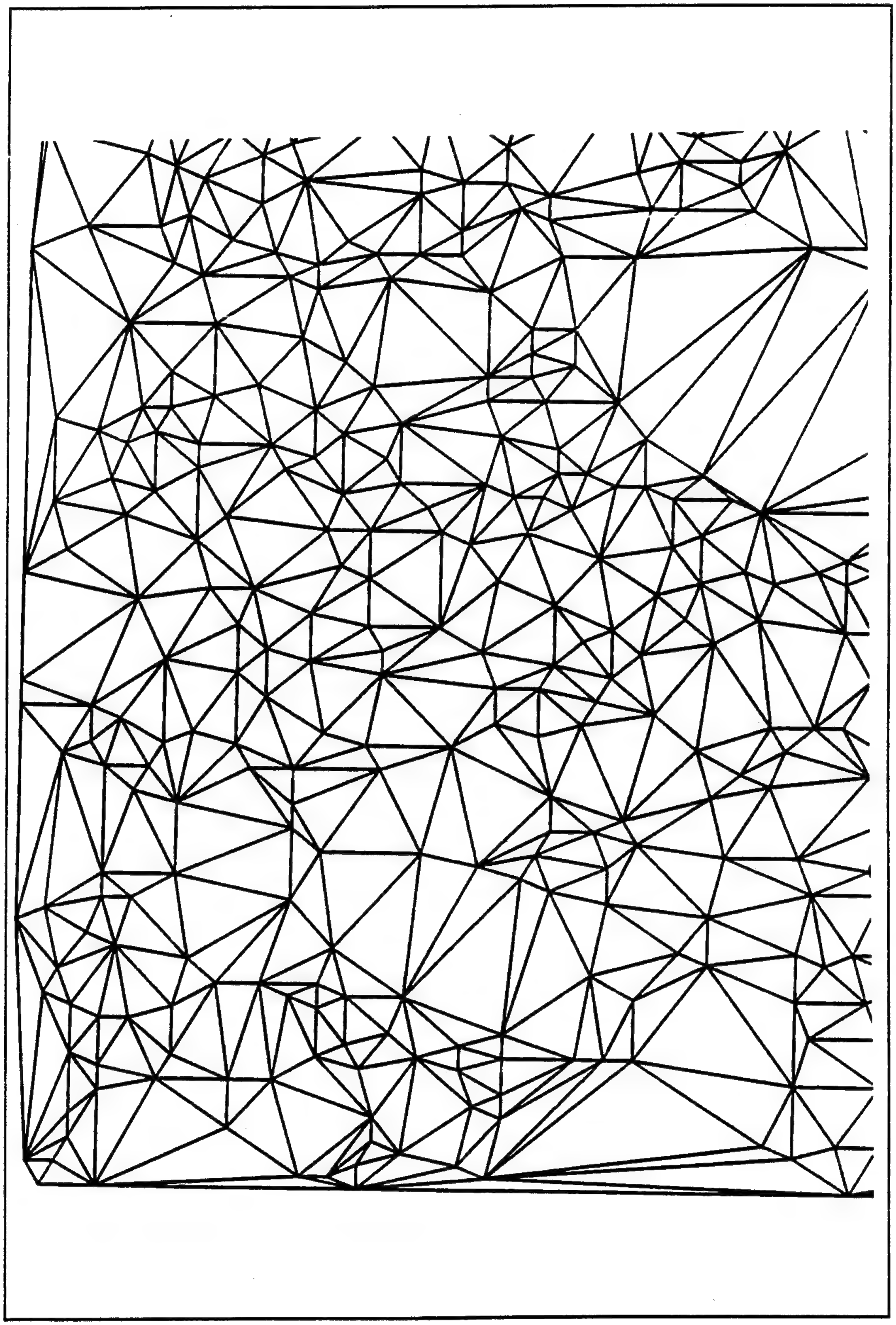


Figure B18. Long Creek TIN - Filter Threshold 7

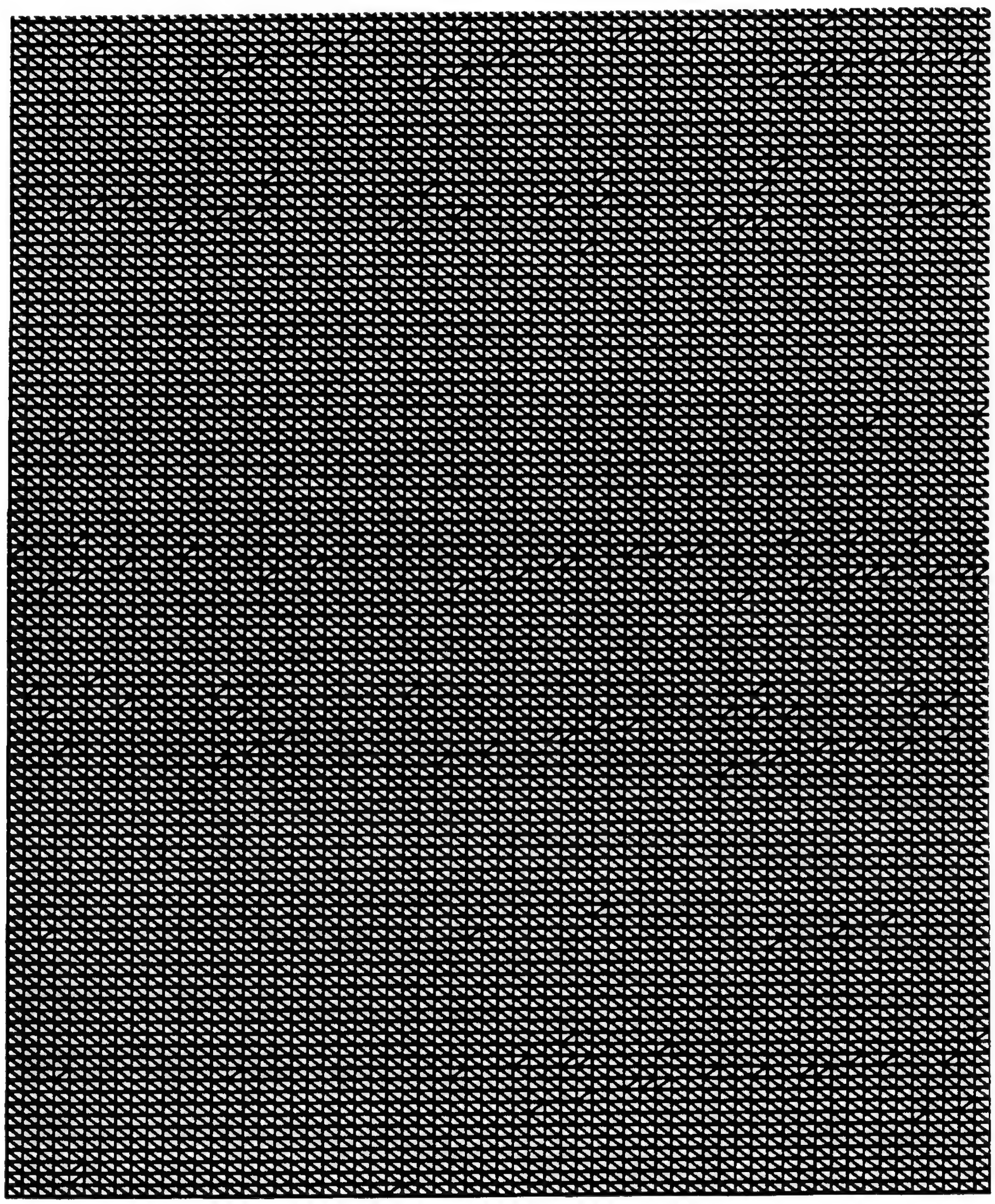


Figure B19.  Beaverdam Creek TIN - No Filtering

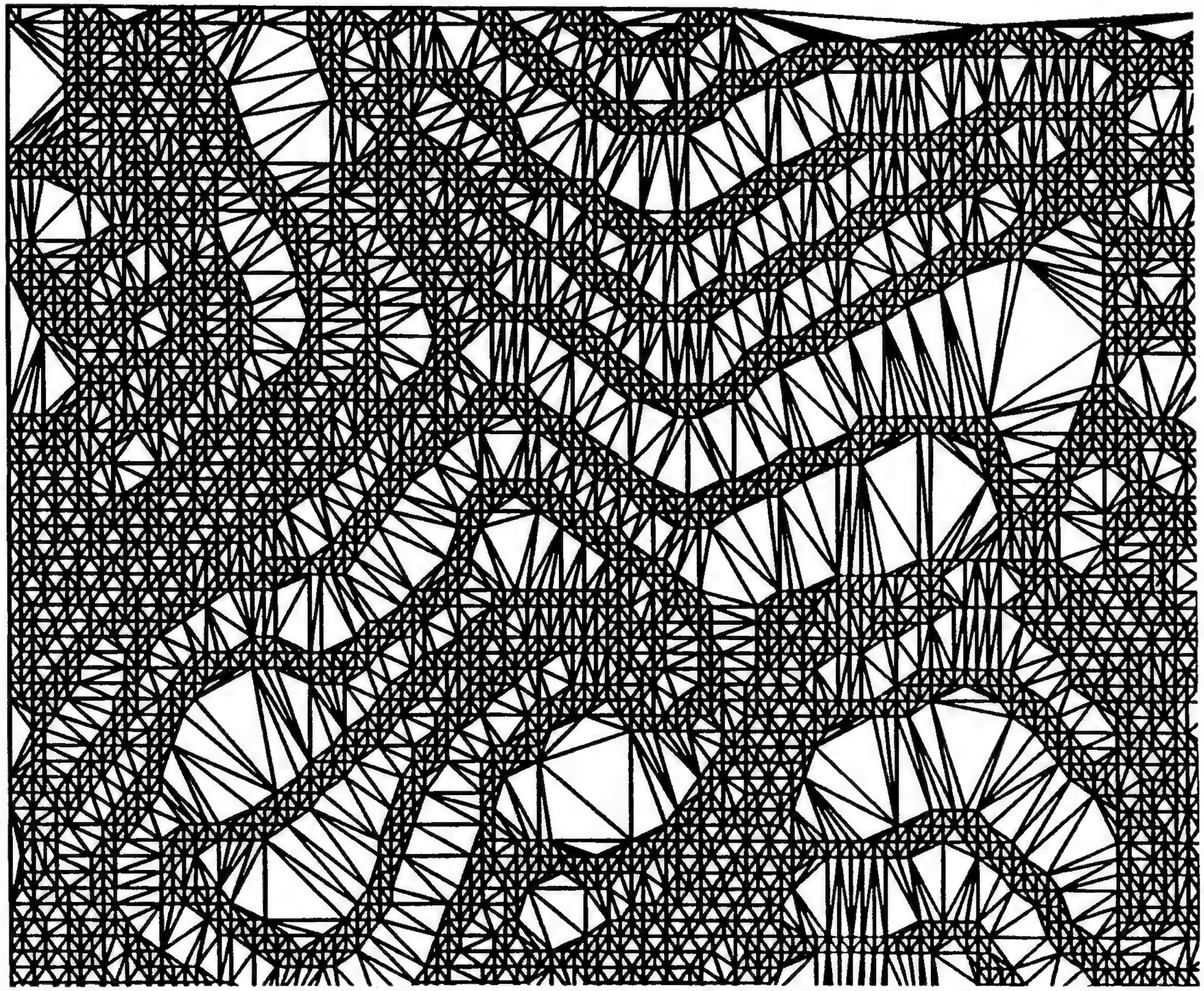


Figure B20. Beaverdam Creek TIN - Filter Threshold 0

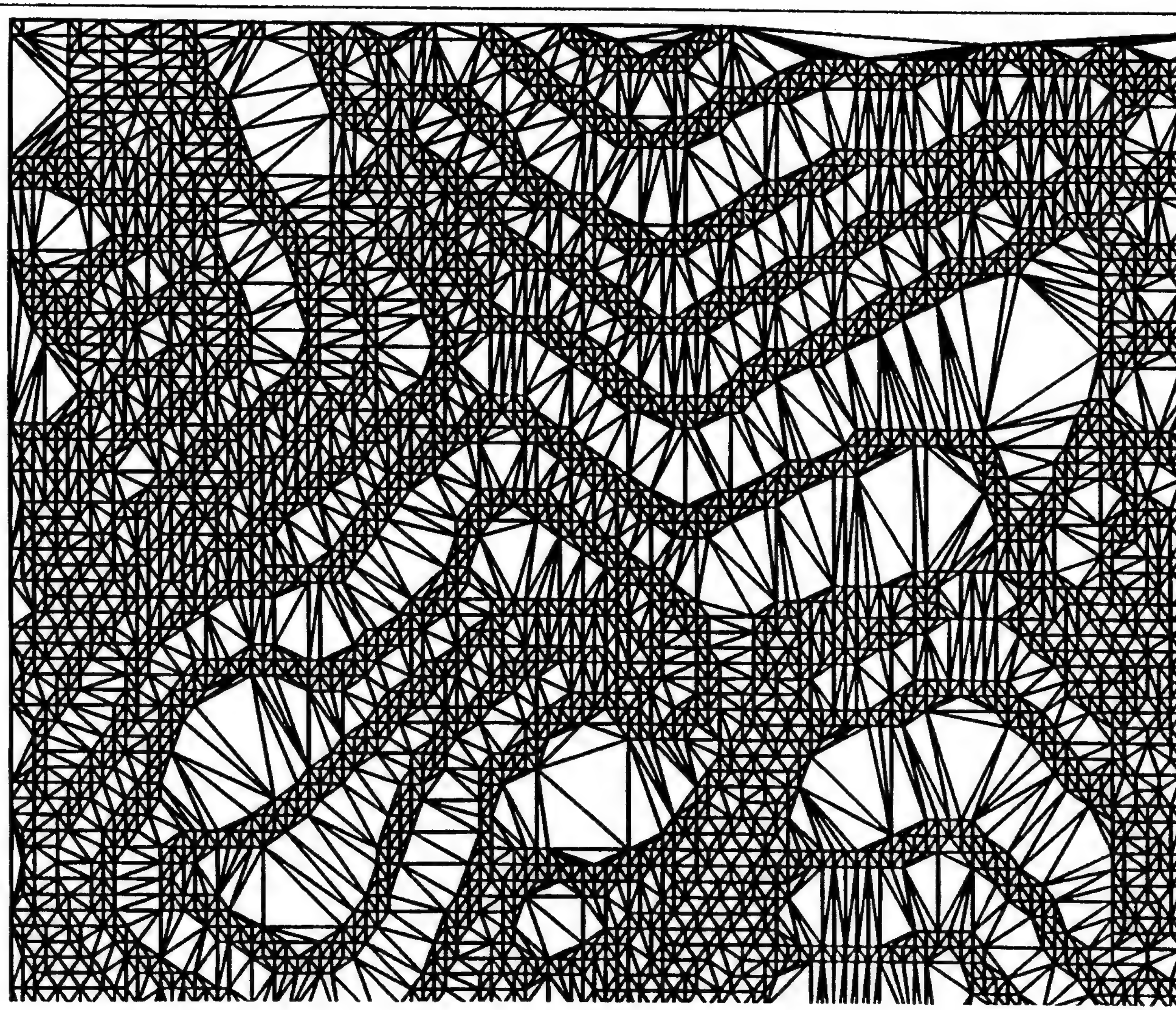


Figure B21. Beaverdam Creek TIN - Filter Threshold 1

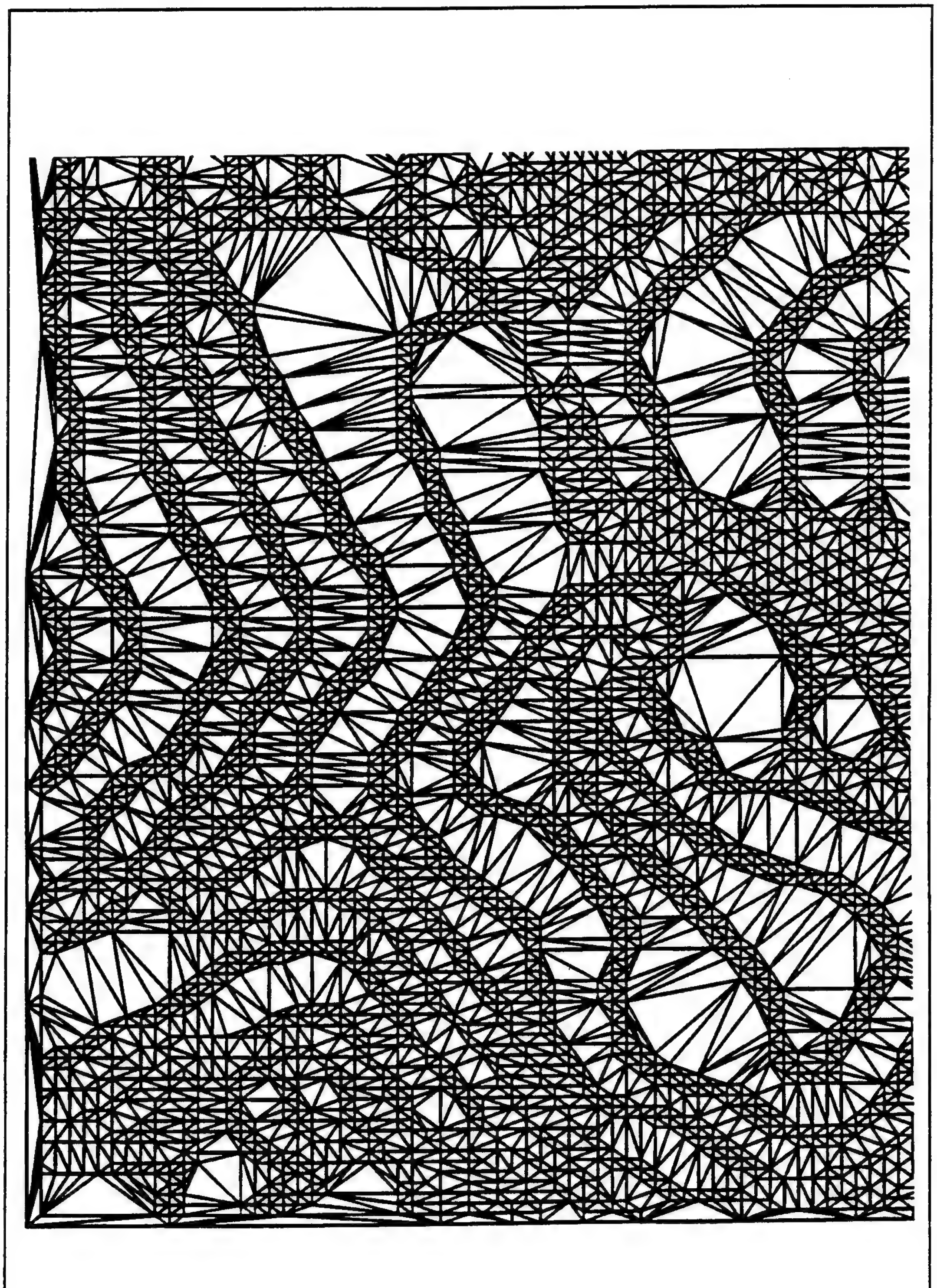


Figure B22. Beaverdam Creek TIN - Filter Threshold 2

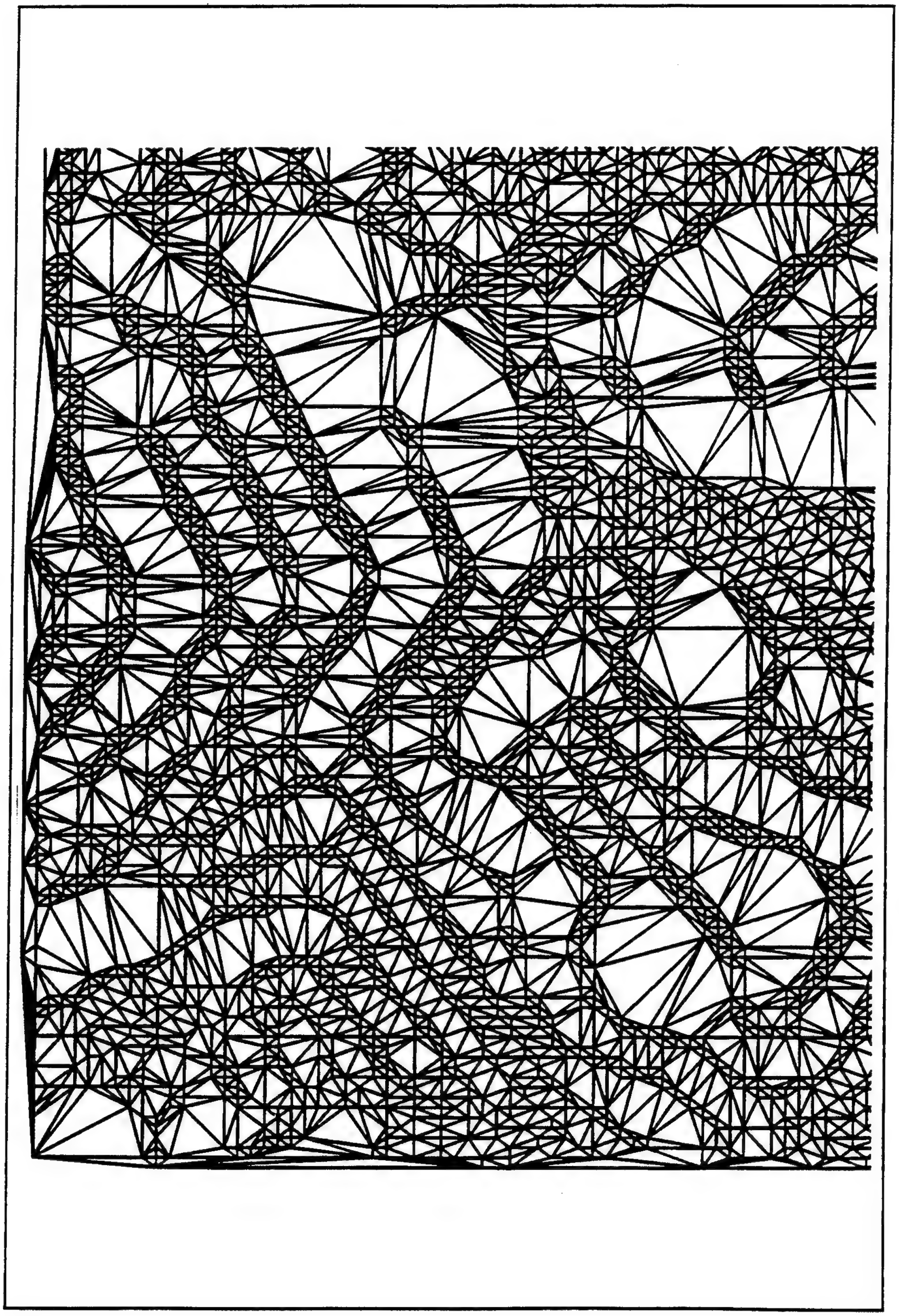


Figure B23. Beaverdam Creek TIN - Filter Threshold 3

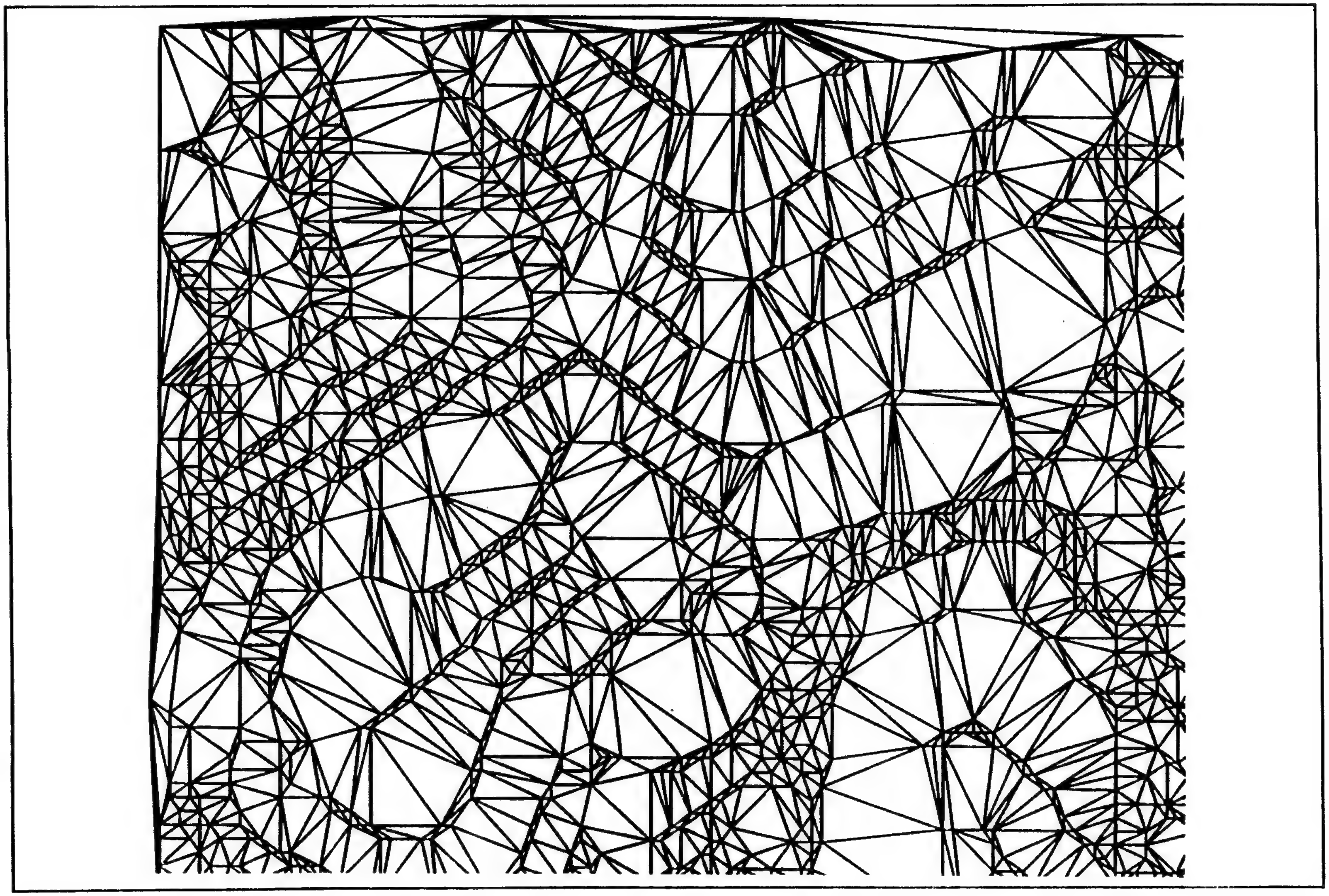


Figure B24.  Beaverdam Creek TIN - Filter Threshold 4

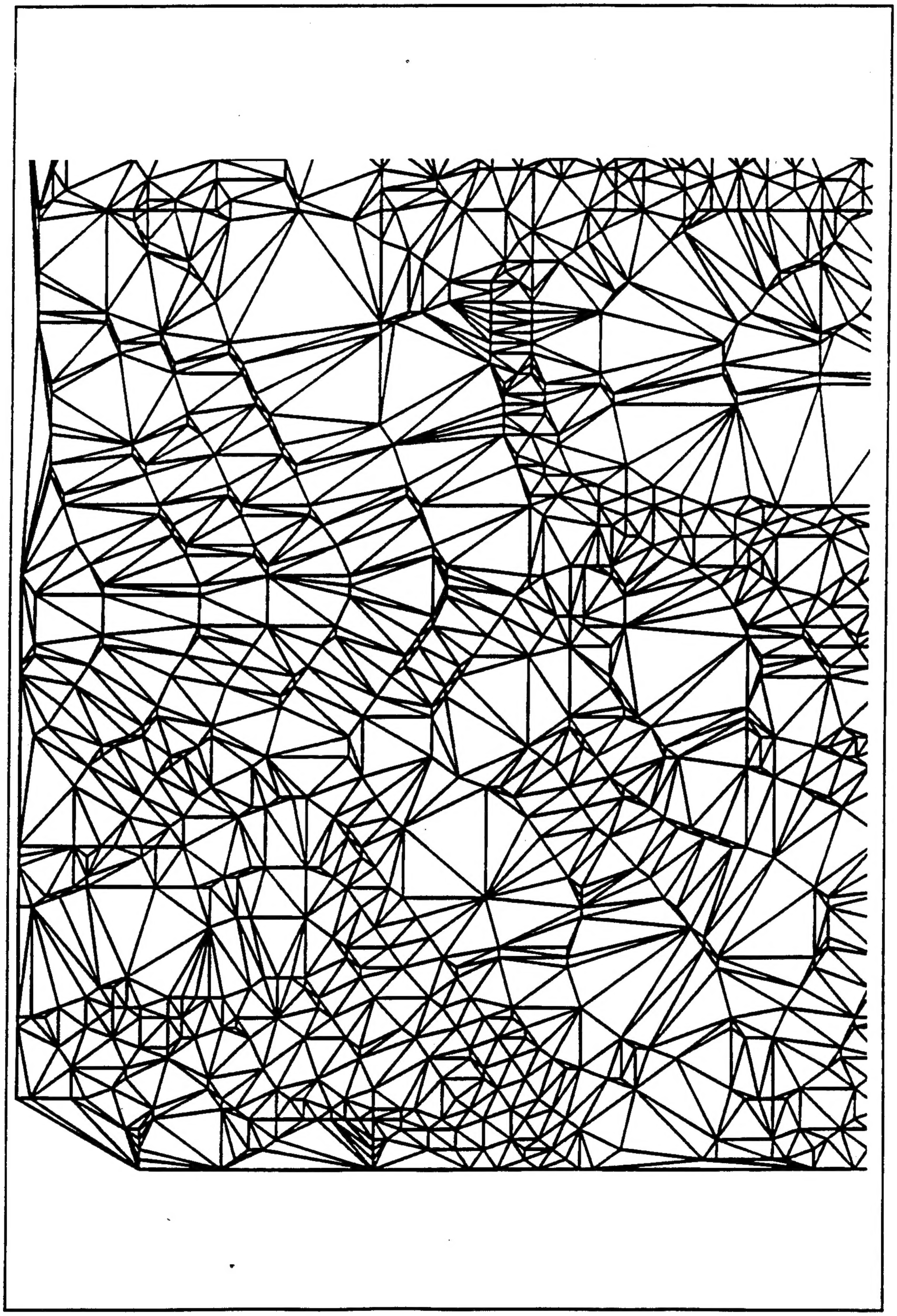


Figure B25. Beaverdam Creek TIN - Filter Threshold 5

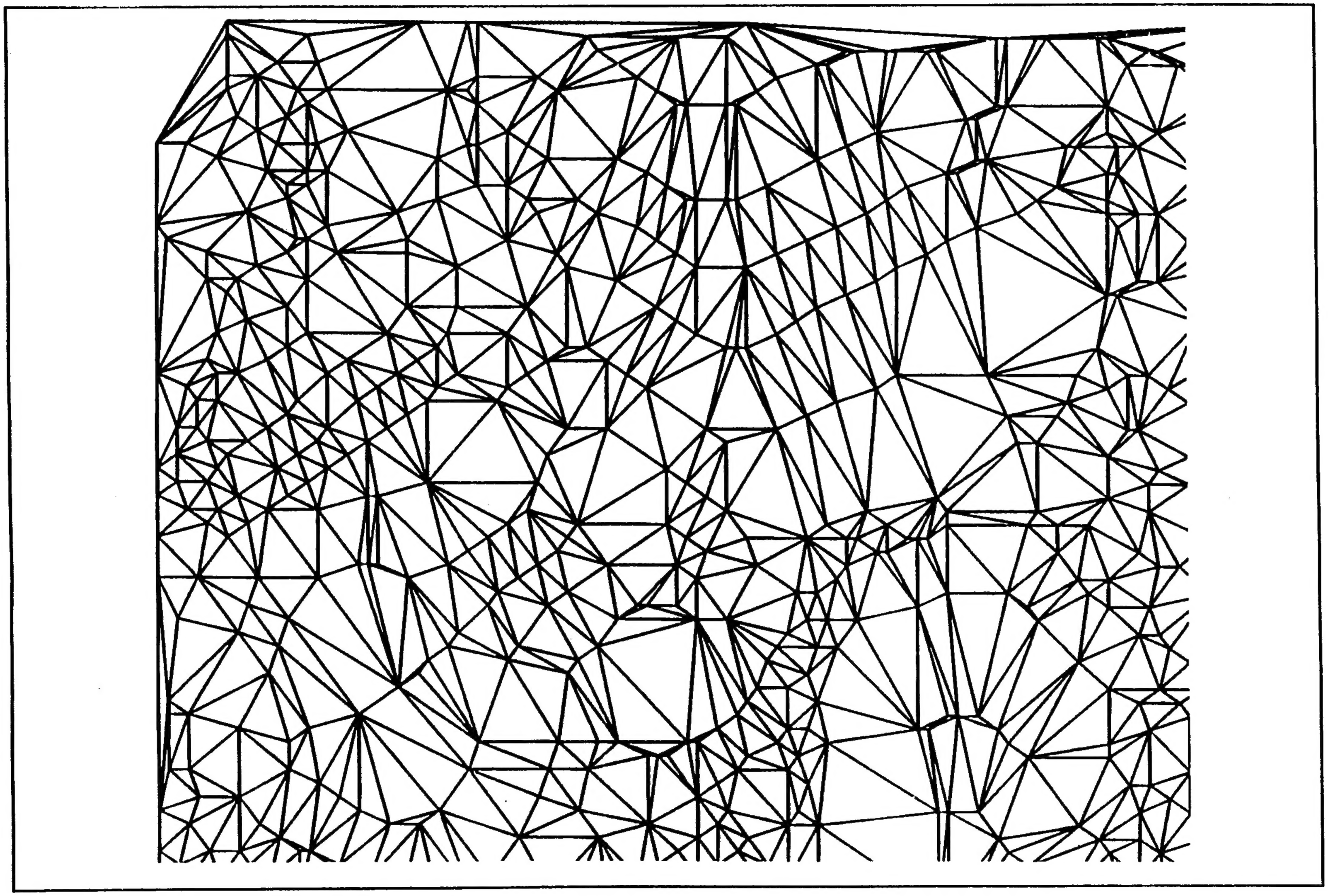


Figure B26. Beaverdam Creek TIN - Filter Threshold 6

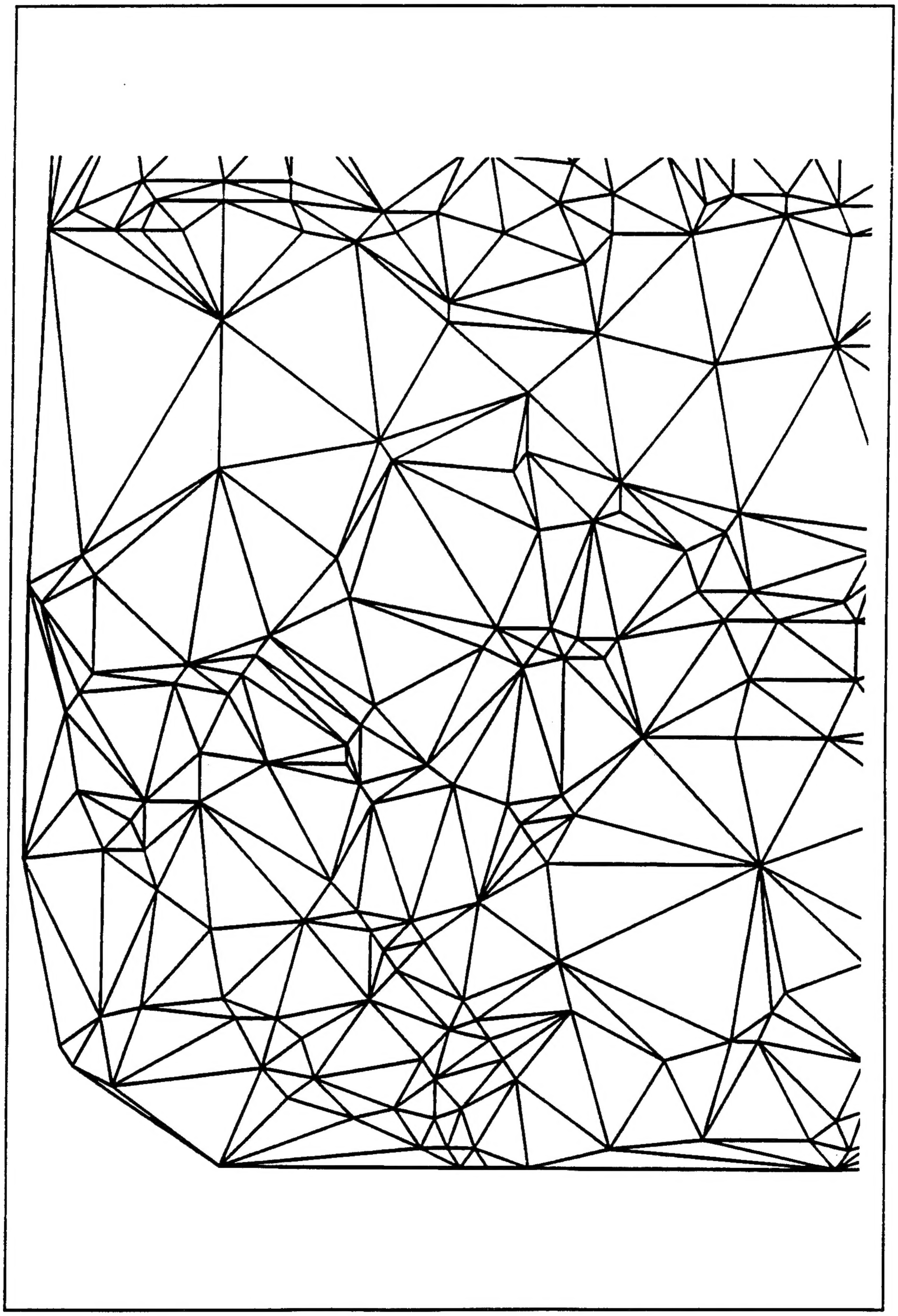


Figure B27. Beaverdam Creek TIN - Filter Threshold 7

# REPORT DOCUMENTATION PAGE

*Form Approved*
*OMB No. 0704-0188*



**1. AGENCY USE ONLY *(Leave blank)***

**2. REPORT DATE**
September 1995

**3. REPORT TYPE AND DATES COVERED**
Final report

**4. TITLE AND SUBTITLE**
Guidance on Generating Hydrologic Model Input from a Triangulated Irregular Network (TIN) Using a Curvature Based Filtering Technique to Extract TIN Vertices from Gridded Digital Elevation Models

**5. FUNDING NUMBERS**

**6. AUTHOR(S)**
Jeffrey D. Jorgeson

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
U.S. Army Engineer Waterways Experiment Station
3909 Halls Ferry Road, Vicksburg, MS 39180-6199

**8. PERFORMING ORGANIZATION REPORT NUMBER**
Technical Report HL-95-7

**9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)**
U.S. Army Corps of Engineers
Washington, DC 20314-1000

**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**
Available from National Technical Information Service, 5285 Port Royal Road, Springfield, VA 22161.

**12a. DISTRIBUTION/AVAILABILITY STATEMENT**

Approved for public release; distribution is unlimited.

**12b. DISTRIBUTION CODE**

**13. ABSTRACT *(Maximum 200 words)***

Extracting input data for hydrologic modeling from triangulated irregular networks (TIN's) is studied. A technique is analyzed that selects points from gridded elevation data for TIN vertices. Elevation points in the grid are filtered out based upon a user defined threshold, with data removal ranging between 20 and 95 percent. Guidance is provided on appropriate filtering thresholds when input data for hydrologic modeling are extracted from the TIN's. The following procedure was used:

*a.* Gridded elevation data for three watersheds representing different terrains were filtered at several thresholds and TIN's were generated from the filtered data.

*b.* Watershed geometric parameters were computed from the TIN's, and parameter differences vrsus the amount of data filtered were determined.

*c.* The geometric parameters were used as input for HEC-1, and differences in hydrographs versus the amount of data fitered were documented.

Removal of approximately 80 percent of the elevation grid points showed no significant changes in discharge hydrographs.

**14. SUBJECT TERMS**
Digital elevation model (DEM)
Digital terrain modeling
Hydrologic modeling
Hydrology
Triangulated irregular network (TIN)
Triangulation

**15. NUMBER OF PAGES**
214

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT | 18. SECURITY CLASSIFICATION OF THIS PAGE | 19. SECURITY CLASSIFICATION OF ABSTRACT | 20. LIMITATION OF ABSTRACT |
|---|---|---|---|
| UNCLASSIFIED | UNCLASSIFIED | | |

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-18
298-102